On Compound Variables and the Asymptotic Behaviour of Their Tails

Master's thesis

Tuuli Tolvanen

University Of Helsinki
Department of Mathematics and Statistics
June 5, 2020

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Faculty of Science	Mathematics and Statistics

Tekijä — Författare — Author

Tuuli Tolvanen

Työn nimi — Arbetets titel — Title

On Compound Variables and the Asymptotic Behaviour of Their Tails

Oppiaine — Läroämne — Subject

Mathematics

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master's thesis	June 2020	49 pages

Tiivistelmä — Referat — Abstract

The objective of this thesis is to introduce the concept of compound variables and explain their use in one application specifically, as the total claim amount of an insurance company can be viewed as a compound variable. We study both the average behaviour as well as the tail behaviour of compound variables.

Before delving into the results concerning the tails of compound variables, we aim to present an overview about the general theory and treat the average behaviour of compound variables first.

We familiarize the reader with rudimentary concepts such as moment and cumulant generating functions. Along the way, the reader will also gain an understanding of both mixed variables as well as compound mixed variables. We state and prove some fundamental results concerning the expectation, variance and moment generating functions of compound variables.

When the concept of compound variable is used to interpret the total claim amount, we also find the number of claims to be of interest. Since it is a random variable, we wish to be able to model it somehow. In the case of a general compound variable, the number of claims simply corresponds to the number of summands in the variable.

We consider compound Poisson variables as a special case of compound variables. The reason for this is that if the counting variable or the number of claims variable is Poisson distributed, then the compound variable is a compound Poisson random variable. We also enhance the modelling of the number of claims by presenting mixing variables into the model.

As a more general version for determining the expectation of a random sum we prove Wald's identity. It does not assume the independence of the counting variable and the increments in the same way we do in the definition of a compound variable.

Towards the end, we shift the focus from general theory and average behaviour to tail behaviour of compound variables. We introduce the reader to the necessary classes of heavy-tailed and subexponential distributions to be able to formulate a few results that give an asymptotically equivalent approximation for the tail function of the compound variable. We prove the result for the case of the negative expectation of the increments (summands). We also present results for the case of non-negative expectation of the increments. Such a situation would be of interest in particular for total claim amounts, if we assume the claims being non-negative random variables.

Avainsanat — Nyckelord — Keywords

Heavy-tails, Compound variables, Asymptotics, Wald's identity

Säilytyspaikka — Förvaringsställe — Where deposited

Kumpula Science Library

Muita tietoja — Övriga uppgifter — Additional information

Contents

Notations and conventions

To begin, we shall introduce some notation and elementary definitions. These conventions are used throughout the thesis unless stated otherwise. The notations below follow the one of [Foss et al., 2011].

- We write $\mathbb{R}^+$ for the non-negative real half-line $[0, \infty)$.
- We write $\mathbb{N}^+$ for strictly positive natural numbers $\{1, 2, 3, \dots\}$.
- By notation $o(g(x))$, we mean: assuming that g is positive, we say that $f(x) = o(g(x))$, if $f(x)/g(x) \to 0$, as x tends to infinity.
- On the other hand, by $O(g(x))$ we denote the following: assuming again the positivity of g, we write $f(x) = O(g(x))$, if $\limsup_{x\to\infty} |f(x)|/g(x) < \infty$.
- By writing $F * G$ we comprehend that it means the convolution of two distributions F and G, that is, the distribution of the random variable $\xi + \eta$, where ξ and η are independent and have distributions F and G, respectively.
- When writing $f(x) \sim g(x)$, we mean that $f(x)/g(x) \to 1$, as $x \to \infty$, given that g is positive. In what follows, x tends to infinity, even if not stated explicitly.
- F^{*n} denotes the n-fold convolution of the distribution F with itself.
- When considering identically distributed random variables $\xi_1, \xi_2, \xi_3, \dots,$ we may write ξ to represent a general ξ_i, $i = 1, 2, 3, \dots$.

Chapter 1

Introduction

In this thesis we study both the average behaviour as well as the tail behaviour of compound variables. For instance, the total claim amount of an insurance company during a year can be viewed as a compound variable.

Before delving into the result concerning the tails of compound variables, we aim to present an overview about the general theory and treat the average behaviour of compound variables first.

So, we consider some relevant theory of compound variables and their average behaviour and make an effort to familiarize the reader with an important concept of compound variables amongst others. We acquaint the reader with concepts such as moment and cumulant generating functions. Along the way, the reader will also gain an understanding of both mixed variables as well as compound mixed variables.

When the concept of compound variable is used to interpret the total claim amount, we also find the number of claims to be of interest. Since it is a random variable, we wish to be able to model it somehow. In the case of a general compound variable, the number of claims simply corresponds to the number of summands in the variable.

We consider compound Poisson variables as a special case of compound variables. The reason for this is that if the counting variable (the number of claims variable) is Poisson distributed, then the compound variable is a compound Poisson distributed random variable. We also enhance the modelling of the number of claims by introducing mixing variables in the model.

Finally, we shift the focus from general theory and average behaviour to tail behaviour of compound variables. We introduce the reader to the necessary classes of heavy-tailed and subexponential distributions to be able to formulate the following results. We have results that give us asymptotic equivalences for the tails of compound variables. We prove the result in the case of the negative expectation of the increments (summands) in more detail.

Chapter 2

Preliminaries

2.1 Necessary definitions

We start by introducing some important definitions. The following fundamental definitions are adapted from [Izyurov, 2019].

Definition 2.1.1. Suppose Ω is a set, and $\mathcal{F}$ is a collection of its subsets. The collection $\mathcal{F}$ is a $\sigma - algebra$, if the following three conditions are satisfied:

(i) $\emptyset \in \mathcal{F}$.

(ii) If $A \in \mathcal{F}$, then $A^c = \Omega \setminus A \in \mathcal{F}$.

(iii) If $A_1, A_2, \ldots$ is a sequence of subsets of Ω such that $A_i \in \Omega$ for $i = 1, 2, \ldots$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition 2.1.2. Given a measurable space $(\Omega, \mathcal{F})$, a function $\mu\colon \mathcal{F} \to \mathbb{R}^+ \cup \{\infty\}$ is called a *measure* if it satisfies the following two properties:

(i) $\mu(\emptyset) = 0$.

(ii) (Countable additivity) If $A_1, A_2, \ldots$ is a sequence of disjoint sets such that $A_i \in \mathcal{F}$ for $i = 1, 2, \ldots$, then

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

A measure is called a *probability measure*, if $\mu(\Omega) = 1$.

Definition 2.1.3. A *probability space* is a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a set, $\mathcal{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.

Definition 2.1.4. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a measurable map ξ from Ω to a measurable space $(\Omega', \mathcal{F}')$ is called a *random variable* (with values in Ω').

Remark. This thesis only considers real-valued random variables.

The definitions we introduce next follow the presentation of [Lehtomaa, 2019].

Definition 2.1.5. A *distribution* P of the random variable ξ is the probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, such that

$$P(B) = \mathbb{P}(\xi^{-1}(B)) = \mathbb{P}(\omega \in \Omega \mid \xi(\omega) \in B)$$

for every $B \in \mathcal{B}(\mathbb{R})$, where $\mathcal{B}(\mathbb{R})$ is the *Borel sigma-algebra of* $\mathbb{R}$.

Definition 2.1.6. A *distribution function* $F \colon \mathbb{R} \to \mathbb{R}$ of a random variable ξ is defined by

$$F(x) = \mathbb{P}(\xi \leq x) = P((-\infty, x]).$$

The distribution and distribution function have one-to-one correspondence, meaning that one determines the other uniquely. Due to this one-to-one correspondence it is common to abuse the terminology and talk about for example distribution when the distribution function is in question.

Definition 2.1.7. The *nth origin moment* a_n of ξ is defined by

$$a_n = \mathbb{E}(\xi^n) = \int_{-\infty}^{-\infty} x^n \mathrm{d}F(x),$$

if $\mathbb{E}(|\xi^n|) < \infty$, $n = 1, 2, \ldots$.

In the above definition F is the distribution function of ξ. Note that $a_1 = \mathbb{E}(\xi)$. For the nth central moment, we make the following definition.

Definition 2.1.8. The *nth central moment* μ_n of ξ is defined by

$$\mu_n = \mathbb{E}\left[(\xi - a_1)^n\right],$$

where $n \geq 2$.

We notice that for the variance σ_ξ^2 of ξ it holds that

$$\sigma_\xi^2 = \mathbb{E}\left[(\xi - \mathbb{E}(\xi))^2\right] = \mu_2.$$

Furthermore, the standard deviation σ_ξ of ξ is defined to be the square root of the variance, that is

$$\sigma_\xi = \sqrt{\mu_2}.$$

The skewness γ_ξ of ξ, which measures the asymmetry of the distribution of ξ about its expectation a_1, is defined as

$$\gamma_\xi = \frac{\mathbb{E}\left[(\xi - a_1)^3\right]}{\sigma^3} = \frac{\mu_3}{\sigma^3}.$$

2.2 Moment and cumulant generating functions

The text [Lehtomaa, 2019] is used as a reference here.

We move on to presenting moment generating functions, which are closely related to calculating moments; the nth origin moment a_n is the nth derivative of the moment generating function evaluated at zero. The concept of moment generating functions is needed when we talk about heavy-tailed distributions in the following chapters.

Definition 2.2.1. *The moment generating function* of ξ is denoted by $M_\xi \colon \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ and defined by

$$M_\xi(s) = \mathbb{E}(e^{s\xi}).$$

If $\mathbb{E}(e^{s\xi}) = \infty$, we say that the moment generating function does not exist at point s.

Remark. It is often said that the moment generating function does not exist when $\mathbb{E}(e^{s\xi}) = \infty$. Actually, what is meant is that it does not exist as a finite number.

We need moment generating functions to be able to talk about cumulant generating functions.

Definition 2.2.2. *The cumulant generating function* of ξ is denoted by $c_\xi \colon \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ and defined to be the natural logarithm of the moment generating function, that is,

$$c_\xi(s) = \ln(M_\xi(s)),$$

If $M_\xi(s) = \infty$, we say that the cumulant generating function does not exist at point s.

Note, that the moment generating function, and therefore also the cumulant generating function, does not necessarily exist at a single point, or at any positive point. On the other hand, the characteristic function of a scalar random variable always exists. By the characteristic function we mean $\mathbb{E}(e^{is\xi})$, where $s \in \mathbb{R}$ and i is the imaginary unit. We shall not further discuss the characteristic function due to its complex nature.

It is worth noting that the moment generating functions exist only for the light-tailed distributions, providing that the argument is in the domain of the function. What we mean by the domain here is the set of points where M_ξ is finite. See Definition 3.0.1 for random variables with heavy-tailed and light-tailed distributions. The moment generating function is log-convex (since the cumulant generating functions are convex), and therefore also convex. So after some point the moment generating function is increasing and once it obtains value ∞, it does not go back to finite values.

Next we provide an optional characterization of the distribution in the form of a well-known result.

Theorem 2.2.3. *If moment generating functions of two random variables coincide and are finite in $(-r, r)$, $r > 0$, then the distributions of the random variables coincide.*

Proof. A version of the proof can be found for example in [Gut, 2005, Chapter 4], where the moment generating functions are treated. □

Remark. The formulation of the theorem in [Gut, 2005] differs a little from ours, but combined with the definition of the moment generating function of the book, it yields the same formulation.

Theorem 2.2.3 means that the moment generating function is an alternative way to specify the distribution of a random variable and that the moment generating function uniquely determines the distribution when it is finite in the neighbourhood of the origin.

One of the properties of the moment generating functions is that they have the derivatives of all orders in the interior of their domain. Observe that 0 is always in the domain since $M_\xi(0) = \mathbb{E}(e^{0\cdot\xi}) = \mathbb{E}(1) = 1 < \infty$. Suppose that s is in the interior of the domain of M_ξ. Then for the nth derivative $M_\xi^{(n)}(s)$ we have that

$$M_\xi^{(n)}(s) = \mathbb{E}(\xi^n e^{s\xi}).$$

In particular, when 0 is in the interior of the domain, that is, M_ξ is finite in a neighbourhood of the origin, then it holds that

$$M_\xi^{(n)}(0) = \mathbb{E}(\xi^n).$$

This is the connection between moment generating functions and calculating moments which we mentioned earlier.

To take a look at the derivatives of cumulant generating functions, for the first order derivative, we have that

$$c'_\xi(s) = \frac{M'_\xi(s)}{M_\xi(s)},$$

and for the second order

$$c^{(2)}_\xi(s) = \frac{M_\xi(s) \cdot M^{(2)}_\xi(s) - (M'_\xi(s))^2}{(M_\xi(s))^2} = \frac{M^{(2)}_\xi(s)}{M_\xi(s)} - \left(\frac{M'_\xi(s)}{M_\xi(s)}\right)^2. \tag{2.1}$$

Therefore

$$c'_\xi(0) = \frac{M'_\xi(0)}{M_\xi(0)} = \frac{\mathbb{E}(\xi e^{0\cdot\xi})}{\mathbb{E}(e^{0\cdot\xi})} = \mathbb{E}(\xi)$$

as well as

$$\begin{aligned} c^{(2)}_\xi(0) &= \frac{M^{(2)}_\xi(0)}{M_\xi(0)} - (c'(0))^2 = \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2 = \mathbb{E}(\xi^2) - 2(\mathbb{E}(\xi))^2 + (\mathbb{E}(\xi))^2 \\ &= \mathbb{E}\left[\xi^2 - 2\xi\mathbb{E}(\xi) + (\mathbb{E}(\xi))^2\right] = \mathbb{E}\left[(\xi - \mathbb{E}(\xi))^2\right] = \sigma^2_\xi, \end{aligned}$$

since

$$-2(\mathbb{E}(\xi))^2 = -2\mathbb{E}(\xi)\mathbb{E}(\xi) = -2\mathbb{E}(\xi\mathbb{E}(\xi)) = \mathbb{E}(-2\xi\mathbb{E}(\xi)).$$

For the derivative of the third order of c_ξ we take the derivative of (2.1) and make the following calculation:

$$\begin{aligned} c^{(3)}_\xi &= \frac{M_\xi \cdot M^{(3)}_\xi - M^{(2)}_\xi \cdot M'_\xi}{(M_\xi)^2} - \frac{(M_\xi)^2 \cdot 2M^{(2)}_\xi \cdot M'_\xi - (M'_\xi)^2 \cdot 2M_\xi \cdot M'_\xi}{(M_\xi)^4} \\ &= \frac{M^{(3)}_\xi}{M_\xi} - \frac{M^{(2)}_\xi \cdot M'_\xi}{(M_\xi)^2} - \frac{2M^{(2)}_\xi \cdot M'_\xi}{(M_\xi)^2} + 2\left(\frac{M'_\xi}{M_\xi}\right)^3 \\ &= \frac{M^{(3)}_\xi}{M_\xi} - \frac{3M^{(2)}_\xi \cdot M'_\xi}{(M_\xi)^2} + 2\left(\frac{M'_\xi}{M_\xi}\right)^3. \end{aligned} \tag{2.2}$$

We left out the arguments of the functions for the sake of clarity. However, all functions are with respect to s. Let us now determine what is the third order derivative of the cumulant generating function of ξ evaluated at 0.

Evaluated at point $s = 0$, formula (2.2) becomes

$$\begin{aligned}
c_\xi^{(3)}(0) &= \mathbb{E}(\xi^3) - 3\mathbb{E}(\xi^2)\mathbb{E}(\xi) + 2(\mathbb{E}(\xi))^3 \\
&= \mathbb{E}(\xi^3) - \mathbb{E}(\xi^2)\mathbb{E}(\xi) - 2\mathbb{E}(\xi^2)\mathbb{E}(\xi) + 2(\mathbb{E}(\xi))^3 + (\mathbb{E}(\xi))^3 - (\mathbb{E}(\xi))^3 \\
&= \mathbb{E}\left[\xi^3 - \xi^2\mathbb{E}(\xi) - 2\xi^2\mathbb{E}(\xi) + 2\xi(\mathbb{E}(\xi))^2 + \xi(\mathbb{E}(\xi))^2 - (\mathbb{E}(\xi))^3\right] \\
&= \mathbb{E}\left[(\xi^2 - 2\xi\mathbb{E}(\xi) + (\mathbb{E}(\xi))^2)(\xi - \mathbb{E}(\xi))\right] \\
&= \mathbb{E}\left[(\xi - \mathbb{E}(\xi))^3\right] \\
&= \gamma_\xi \cdot \sigma^3.
\end{aligned}$$

So we have seen that for the second and third order derivatives evaluated at 0 it holds that $c_\xi^{(n)}(0) = \mathbb{E}\left[(\xi - \mathbb{E}(\xi))^n\right] = \mu_n$, whenever $n = 2, 3$.

2.3 Compound variables and total claim amounts

This section is based on [Daykin et al., 1994, Chapter 3] and [Lehtomaa, 2019].

We will begin by first defining the important concept of counting variables. They have a significant role when it comes to compound variables.

Definition 2.3.1. A random variable τ is called a *counting variable* if

$$\mathbb{P}(\tau \in \{0, 1, 2, \dots\}) = 1.$$

Now we go on to defining the single most important concept of this thesis.

Definition 2.3.2. Let F be a distribution function and τ a counting variable. The random variable $S_\tau = \xi_1 + \cdots + \xi_\tau$ is called a *compound variable with parameter (τ,F)*, if the following two conditions are satisfied:

(i) $\tau, \xi_1, \xi_2, \dots$ are independent.

(ii) The distribution function of $\xi_1, \xi_2, \dots$ is F.

At this point we make a comment that conditions (i) and (ii) are seldom satisfied in practice.

A concrete example of compound variables in this thesis is the total claim amount. Next we will clarify how we interpret the situation. In the case of an accident, the policy-holder makes a claim informing the insurance company of the occurred accident. The insurance company must then compensate the policy-holder for the losses that the accident caused. Typically it is an economic loss in question. It is natural to model the compensation as a non-negative random variable. We call the aforementioned random variable *claim size*. The sum of these random variables is the *total claim amount*. We generally consider the sum of claims within a year in a given insurance portfolio (fixed collection of insurance contracts).

Using the notation of Definition 2.3.2, we could understand that the random variable S_τ symbolizes the total claim amount. Mathematically, S_τ can be interpreted as a *random sum*, since both the values of the summands as well as the number of summands are random. In this situation each summand ξ_i, $i = 1, \ldots, \tau$, an increment, would represent the size of the ith claim. The number of claims occurred during one year (or other given time period) is τ in this set-up.

Even though the definition of a compound variable itself does not restrict the sign of the increments, in the case of total claim amount, it is natural to assume that the increments are non-negative random variables. However, there are no technical reasons as to why the increments could not admit negative values. In fact, as pointed out in [Daykin et al., 1994], the concept of negative claim sizes comes in handy in situations where an event causes the wealth of the insurer to increase. Nonetheless, the regular case is that the claims decrease the wealth of the insurer.

According to [Daykin et al., 1994], when bringing the compound model into practical applications, the division of the portfolio as per to the line of business ought to be made. Each line of business is modelled using its own compound variable. These actions make the assumption about the existence of the common distribution function F of ξ_i's, $i = 1, 2, \ldots$, for each business line reasonable. The reasonability is obtained at least for moderate time periods and when the fluctuations of monetary values are disregarded.

We continue by a theorem about the moment generating function of a compound variable.

Theorem 2.3.3. *Let $S_\tau = \xi_1 + \cdots + \xi_\tau$ be a compound variable with parameter (τ, F), where F is the distribution function of ξ. Let M_τ be the moment generating function of τ and c_ξ the cumulant generating function of ξ. Then*

$$M_{S_\tau}(s) = M_\tau(c_\xi(s)),$$

for all $s \in \mathbb{R}$. We make the agreement that $M_{S_\tau}(s) = \infty$ in case that $c_\xi(s) = \infty$.

Proof. Write

$$M_{S_\tau}(s) = \sum_{t=0}^{\infty} \mathbb{E}(e^{s(\xi_1+\cdots+\xi_t)} \mathbb{1}(\tau = t)) = \sum_{t=0}^{\infty} \mathbb{E}(e^{s(\xi_1+\cdots+\xi_t)}) \mathbb{P}(\tau = t).$$

The second equality holds because of the independence of τ and ξ_i's, $i = 1, 2, 3, \ldots$.

We continue by writing

$$
\begin{aligned}
\sum_{t=0}^{\infty} \mathbb{E}(e^{s(\xi_1+\cdots+\xi_t)})\mathbb{P}(\tau = t) &= \sum_{t=0}^{\infty} M_\xi(s)^t \mathbb{P}(\tau = t) \\
&= \sum_{t=0}^{\infty} e^{\ln(M_\xi(s)^t)} \mathbb{P}(\tau = t) \\
&= \sum_{t=0}^{\infty} e^{t\ln(M_\xi(s))} \mathbb{P}(\tau = t) \\
&= \sum_{t=0}^{\infty} e^{tc_\xi(s)} \mathbb{P}(\tau = t) \\
&= \mathbb{E}(e^{\tau c_\xi(s)}) \\
&= M_\tau(c_\xi(s)).
\end{aligned}
$$

The first equality holds because of the independence of ξ_i's, $i = 1, 2, 3, \ldots$. □

2.3.1 Poisson distribution and compound Poisson variables

In this section also [Daykin et al., 1994, Chapter 2] is used as a reference.

As a special case of compound variables we introduce compound Poisson variables. The advantage of being able to model the number of claims variable as a Poisson distributed one is that they have a lot of useful properties, such as additivity. The sum of independent Poisson variables is also a Poisson variable. The parameter of the sum variable is the sum of the parameters of the summands. In fact, if the parameter λ of a Poisson variable happens to be an integer, the variable can be thought of as a sum of λ pieces of independent Poisson variables, each having parameter 1. Also, the variable can be thought of as a sum of 2λ pieces of independent Poisson variables, each having parameter $0,5$, and so on. This idea naturally leads us to defining infinite divisibility (following the representation of [Embrechts et al., 2008, Section 2.2]).

Definition 2.3.4. A random variable ξ (and its distribution) is *infinite divisible* if and only if we can decompose it in law:

$$\xi \stackrel{d}{=} \xi_{n1} + \cdots + \xi_{nn}$$

for every n, where $\{\xi_{nk}\}_{k=1,\ldots,n}$ are independent, identically distributed random variables with possibly different common distribution for different n.

Intuitively, we can think that for example the original Poisson parameter λ is divided into smaller parts, say of size α. Amount α is the Poisson parameter for the summands. Take the number of summands required to form the Poisson(λ) variable to be x, so that $x\alpha = \lambda$. This action can be performed infinitely many times, as there are infinitely many combinations of α and x that yield λ when multiplied. It is therefore clear that the Poisson distribution is an infinitely divisible one.

Another upside to Poisson variables lies in the approximation of the distribution function. Although there is an exact formula for calculating the distribution function $F(t)$ of a Poisson variable, the amount of time needed for performing the calculations for large values of t is excessive. Therefore it is convenient that there exist various approximation methods for the distribution function of a Poisson variable, which give acceptable levels of accuracy.

Furthermore, when the number of claims variable τ has a Poisson distribution, then the total claim amount S_τ is a compound Poisson variable. Hence we can transfer some of the theory of the Poisson variables to the compound case as well.

To begin, we characterize Poisson distributed random variables by defining their probability mass function.

Definition 2.3.5. We say that a random variable τ has a *Poisson distribution with parameter* $\lambda > 0$, if

$$\mathbb{P}(\tau = t) = e^{-\lambda}\frac{\lambda^t}{t!},$$

where $t = 0, 1, 2, \ldots$.

The Poisson distribution is a discrete probability distribution. An example of an application would be to model the amount of cars passing by a certain point on the road during a time interval $(0, t)$ as a Poisson($t\lambda$) distributed random variable. This example was mentioned in the exercise section of [Tuominen, 2010]. Another application that is of main interest in this thesis is of course the number of claims variable being Poisson distributed. The parameter λ could be seen as the *rate* or number of events happening during one time unit. Hence it is also natural that the Poisson parameter is the expectation of the distribution, as we will see in Theorem 2.3.8.

The following lemma shows that the Poisson distribution is the limit of the binomial distribution. The number of trials n is assumed to be large and the success probability p_n is assumed to be small. Therefore the Poisson distribution can be thought of as a suitable distribution for modelling the number of occurrences of a rare event in a large population, as was mentioned in [Tuominen, 2010, Section 2]. The Poisson distribution might be a more desirable option for modelling compared to the binomial distribution.

Lemma 2.3.6. *If $X_n \sim Bin(n, p_n)$, that is*

$$\mathbb{P}(X_n = k) = \binom{n}{k} {p_n}^k (1-p_n)^{n-k}, \quad k = 0, 1, 2, \dots, n,$$

and additionally $n \to \infty$ and $p_n \to 0$ as $n \to \infty$ so that $np_n \to \lambda$, then

$$\mathbb{P}(X_n = k) \to e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

as $n \to \infty$. The limit is the probability mass function of a Poisson(λ) distributed random variable.

Proof. Since it was assumed that $np_n \to \lambda$ as $n \to \infty$ and $p_n \to 0$, we can write

$$np_n = \lambda + o(1) \Leftrightarrow p_n = \frac{\lambda + o(1)}{n},$$

where $o(1) \to 0$ as $n \to \infty$. Then we rewrite the probability mass function of a binomial random variable in the following manner:

$$\begin{aligned} \mathbb{P}(X_n = k) &= \binom{n}{k} {p_n}^k (1-p_n)^{n-k} \\ &= \frac{n \cdot (n-1) \cdots 2 \cdot 1}{k!(n-k) \cdot (n-k-1) \cdots 2 \cdot 1} \cdot \left(\frac{\lambda + o(1)}{n}\right)^k \cdot \left(1 - \frac{\lambda + o(1)}{n}\right)^{n-k}. \end{aligned}$$

Since $k \leq n$, the first factor reduces to

$$\frac{n \cdot (n-1) \cdots (n-k+1)}{k!}.$$

Then taking the limit of $\mathbb{P}(X_n = k)$, as $n \to \infty$, we have that it equals

$$\lim_{n\to\infty} \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \cdot \lim_{n\to\infty} \left(\frac{\lambda + o(1)}{n}\right)^k \cdot \lim_{n\to\infty} \left(1 - \frac{\lambda + o(1)}{n}\right)^{n-k}.$$

From the first factor we can take $\frac{1}{k!}$ out of the limit and it is a factor we want to have in the final answer. The numerator stays the same for now. From the second factor we calculate that

$$\lim_{n\to\infty} \big(\lambda + o(1)\big)^k = \lambda^k,$$

and leave the denominator as is. Next we find out the limit for $\left(1-\frac{\lambda+o(1)}{n}\right)^n$:

$$\begin{aligned}\lim_{n\to\infty}\left(1-\frac{\lambda+o(1)}{n}\right)^n &= \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n\\ &= \lim_{n\to\infty} e^{\ln\left((1-\frac{\lambda}{n})^n\right)}\\ &= \lim_{n\to\infty} e^{n\ln(1-\frac{\lambda}{n})}\\ &= e^{\lim_{n\to\infty} n\ln(1-\frac{\lambda}{n})}\\ &= e^{\lim_{n\to\infty}\frac{\ln(1-\frac{\lambda}{n})}{\frac{1}{n}}}.\end{aligned} \tag{2.3}$$

Due to the numerator and denominator both converging to 0, we use L'Hôpital's rule to figure out the limit. Taking the derivatives and making the division, we get

$$\lim_{n\to\infty}\frac{\frac{\lambda}{n^2-\lambda n}}{-\frac{1}{n^2}} = -\lambda.$$

So, (2.3) becomes $e^{-\lambda}$.

All in all, what is left yet undetermined is

$$\lim_{n\to\infty}(n\cdot(n-1)\cdots(n-k+1))\cdot\lim_{n\to\infty}\left(\frac{1}{n}\right)^k\cdot\lim_{n\to\infty}\left(1-\frac{\lambda+o(1)}{n}\right)^{-k}. \tag{2.4}$$

We aim to show that (2.4) equals 1, which concludes the proof. Rewrite (2.4) as follows:

$$\lim_{n\to\infty}\frac{(n\cdot(n-1)\cdots(n-k+1))}{n^k\left(1-\frac{\lambda+o(1)}{n}\right)^k}.$$

The factor in the denominator in the parentheses goes to $1-0=1$ as n approaches infinity. We are left with

$$\lim_{n\to\infty}\frac{(n\cdot(n-1)\cdots(n-k+1))}{n^k} = \lim_{n\to\infty}\left(\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n}\right) = 1.$$

We have therefore shown that (2.4) indeed equals 1 and we finish by concluding that

$$\lim_{n\to\infty}\mathbb{P}(X_n=k) = e^{-\lambda}\frac{\lambda^k}{k!},$$

as desired. $\square$

We move on to determining the moment generating function for Poisson variables. We shall be using this result when treating compound Poisson variables.

Theorem 2.3.7. *Let τ be a Poisson distributed random variable with parameter λ. Then for the moment generating function we have*

$$M_\tau(s) = e^{\lambda(e^s-1)},$$

for all $s \in \mathbb{R}$.

Proof. For the moment generating function, we calculate

$$M_\tau(s) = \mathbb{E}(e^{\tau s}) = \sum_{t=0}^{\infty} \mathbb{P}(\tau = t)e^{ts} = \sum_{t=0}^{\infty} e^{-\lambda}\frac{\lambda^t}{t!}e^{ts} = e^{-\lambda}\sum_{t=0}^{\infty}\frac{(\lambda e^s)^t}{t!} = e^{-\lambda}e^{\lambda e^s} = e^{\lambda(e^s-1)}.$$

For the second to last equality, we noticed that the sum is the Maclaurin series of the exponential function with power λe^s. $\square$

Next, we formulate a result concerning the first central and origin moments of a Poisson distributed random variable.

Theorem 2.3.8. *Let τ be a Poisson distributed random variable with parameter λ. Then for the expectation, variance and skewness we have*

$$\mathbb{E}(\tau) = \mathrm{Var}(\tau) = \lambda, \quad \textit{and}$$
$$\gamma_\tau = \frac{1}{\sqrt{\lambda}}.$$

Proof. For the expectation, we have

$$\mathbb{E}(\tau) = M'_\tau(0) = e^{\lambda e^0}\lambda e^0 e^{-\lambda} = \lambda.$$

The variance is calculated in the following manner:

$$\begin{aligned}
\mathrm{Var}(\tau) &= \mathbb{E}\left[(\tau - \mathbb{E}(\tau))^2\right] \\
&= \mathbb{E}\left[(\tau - \lambda)^2\right] \\
&= \mathbb{E}\left[\tau^2 - 2\tau\lambda + \lambda^2\right] \\
&= \mathbb{E}(\tau^2) - 2\lambda\mathbb{E}(\tau) + \lambda^2 \\
&= \mathbb{E}(\tau^2) - 2\lambda^2 + \lambda^2 \\
&= M_\tau^{(2)}(0) - \lambda^2 \\
&= \left[e^{\lambda e^0}\lambda e^0 + \lambda e^0 e^{\lambda e^0}\lambda e^0\right]e^{-\lambda} - \lambda^2 \\
&= \lambda.
\end{aligned}$$

The skewness is obtained upon noticing that $\mu_3 = \lambda$ and $\sigma^3 = \lambda\sqrt{\lambda}$, and making the division. $\square$

By the previous theorem the expectation and variance are equal, and this fact can be used to evaluate whether a Poisson model is suitable for modelling the claim number variable. The expectation and variance can be calculated from the data, and if they are equal, then the Poisson model might be a good fit. However, if they are not equal, it can be anticipated that the Poisson model will be an insufficient description of the data. An example of simulations made using the probability mass function can be found in [Daykin et al., 1994, Chapter 2]. It is observed that in fact, the Poisson model is not an appropriate one when the expectation and variance are unequal.

We are now ready to define the compound Poisson variable.

Definition 2.3.9. Let S_τ be a compound variable with parameter (τ, F). If τ has a Poisson distribution with parameter λ, then S_τ is called a *compound Poisson variable with parameter* (λ, F).

We formulate a result similar to Theorem 2.3.3 concerning compound Poisson variables.

Corollary 2.3.10. *Let $S_\tau = \xi_1 + \cdots + \xi_\tau$ be a compound Poisson variable with parameter (λ, F), where F is the distribution function of ξ. Let M_ξ be the moment generating function of ξ. Then*

$$M_{S_\tau}(s) = e^{\lambda(M_\xi(s)-1)},$$

for all $s \in \mathbb{R}$. We agree that $M_{S_\tau}(s) = \infty$, if $M_\xi(s) = \infty$.

Proof. By Theorem 2.3.3, for a general compound variable it holds that

$$M_{S_\tau}(s) = M_\tau(c_\xi(s)),$$

and by Theorem 2.3.7,

$$M_\tau(s) = e^{\lambda(e^s-1)}.$$

Therefore

$$M_{S_\tau}(s) = M_\tau(c_\xi(s)) = e^{\lambda(e^{c_\xi(s)}-1)} = e^{\lambda(e^{\ln(M_\xi(s))}-1)} = e^{\lambda(M_\xi(s)-1)}.$$

$\square$

2.3.2 Mixing variables and mixed models

In this section also [Daykin et al., 1994, Chapter 2] is used as a reference.

There is a need to take into account the effect of background factors such as weather conditions in the model. We can alter the Poisson model of the number of claims to be more accurate and applicable. This could be done via introducing an auxiliary variable to the model, turning the model into a mixed Poisson one. The mixed Poisson distribution generalizes the Poisson distribution. In practice it is often a more realistic option for modelling the number of claims in comparison with plain Poisson distribution as it allows a more detailed model of the reality. We will return to the topic of mixing and mixed variables shortly.

By the term risk intensity we mean the intensity rate of the events that cause claims. The need for a mixed Poisson model arises from the volatility of risk intensity. The variations in risk intensity are caused by external background factors. These factors such as weather (mentioned above) and economic conditions have an effect on the intensity of the claims. An example of an economic impact could be that during a recession, the employees' working hours are anticipated to diminish, resulting in fewer work-related accidents. On the contrary, during an economic boom the working hours are likely to be increased, therefore we can expect more accidents.

In case the variations in risk intensity are deterministic, the Poisson model is applicable. Deterministic variations include for example those relating to daytime and nighttime. These variations can be anticipated and there is no significant randomness connected to them. When the variation in claim intensity is random, it is common that the variation can be understood as random changes of the parameter from the expectation of the distribution, which is λ whenever $\tau \sim \text{Poisson}(\lambda)$.

Now we will discuss the aforementioned mixing variable in more detail. A multiplicative factor Q depicts the change of the parameter λ. Let Q be a random variable such that $\mathbb{E}(Q) = 1$. Furthermore, let Q be such that if the claim intensity is at its expected level during a certain time period, then the value q of Q would equal 1. If $q > 1$, then the intensity is higher than expected, and in the case where $0 < q < 1$, the intensity is lower than expected.

Definition 2.3.11. We call the above described random variable Q a *mixing variable* or *structure variable*.

Definition 2.3.12. Let Q be a mixing variable. Let $\lambda > 0$ be a constant. We say that τ is a *mixed Poisson variable* and write $\tau \sim \text{Poisson}(\lambda, Q)$ for such a counting variable that the following holds true:

$$F_{\tau|Q}(t \mid q) = \mathbb{P}(\tau \le t \mid Q = q) = e^{-\lambda q} \sum_{h=0}^{t} \frac{(\lambda q)^h}{h!},$$

where $t = 0, 1, 2, \ldots$.

We make a remark that if $Q = q$ is fixed, then the conditional distribution function of the claim number variable, $\mathbb{P}(\tau \le t \mid Q = q)$, is the distribution function of a Poisson(λq) distributed random variable.

We can understand the value of the mixing variable to be drawn from the distribution of Q, F_Q. The drawing is performed at the beginning of each time period, a year for example. Let q be the drawn value of a certain time period. Then the Poisson parameter for that time period equals λq. To clarify the meaning of mixing variables, the different values of a mixing variable can be thought of as corresponding to each combination of possible outcomes of the surrounding factors. Thus the different values of Q represent the alternative states of the background factors.

We observe that whenever the mixing variable Q is actually a constant 1, then the mixed Poisson variable τ reduces to a regular Poisson variable. This happens because

$$\mathbb{P}(\tau \le t \mid Q = q) = e^{-\lambda q} \sum_{h=0}^{t} \frac{(\lambda q)^h}{h!} = e^{-\lambda} \sum_{h=0}^{t} \frac{\lambda^h}{h!} = \mathbb{P}(\tau \le t),$$

which is the distribution function for Poisson(λ) distribution. This could have been seen also by utilizing what was stated above, that the conditional distribution function of τ is that of a Poisson(λq) distributed random variable, which simplifies to Poisson(λ) distribution, whenever $Q = q = 1$.

We further define the concept of compound mixed Poisson variables.

Definition 2.3.13. If τ has a mixed Poisson distribution with parameter (λ, Q), we call S_τ a *compound mixed Poisson variable with parameter* (λ, Q, F).

We now formulate a result similar to Theorem 2.3.3 for compound mixed Poisson variables.

Corollary 2.3.14. *Let $S_\tau = \xi_1 + \cdots + \xi_\tau$ be a compound mixed Poisson variable with parameter (λ, Q, F), where F is the distribution function of ξ. Let M_Q and M_ξ be the moment generating functions of Q and ξ respectively. Then*

$$M_{S_\tau}(s) = M_Q(\lambda(M_\xi(s) - 1)),$$

for all $s \in \mathbb{R}$. We agree that $M_{S_\tau}(s) = \infty$, if $M_\xi(s) = \infty$.

Proof. By Theorem 2.3.3, for a general compound variable it holds that

$$M_{S_\tau}(s) = M_\tau(c_\xi(s)),$$

and we also have that

$$M_\tau(s) = \mathbb{E}[\mathbb{E}(e^{\tau s} \mid Q)] = \mathbb{E}(e^{\lambda Q(e^s-1)}) = M_Q(\lambda(e^s - 1)).$$

Therefore

$$M_{S_\tau}(s) = M_\tau(c_\xi(s)) = M_Q(\lambda(e^{c_\xi(s)} - 1)) = M_Q(\lambda(e^{\ln M_\xi(s)} - 1)) = M_Q(\lambda(M_\xi(s) - 1)).$$

□

We have hereby determined the moment generating functions for compound variables, Poisson variables, compound Poisson variables as well as compound mixed Poisson variables.

Chapter 3

Heavy-tailed distributions

The references for this chapter are [Denisov et al., 2010] and [Lehtomaa, 2019].

Let us begin with defining the concept that lent its name to this chapter.

Definition 3.0.1. A random variable ξ has a *heavy-tailed* distribution, if

$$M_\xi(s) = \mathbb{E}(e^{s\xi}) = \infty,$$

for all $s > 0$, and a *light-tailed* distribution otherwise.

Remark. By Definition 3.0.1 we actually mean that ξ has a heavy right tail. However, this is often abbreviated and we talk about heavy-tails, especially when it is clear which tail is of interest.

In the case of heavy-tailed distributions, there does not exist a neighbourhood of the origin, where the moment generating function would be finite. Therefore in particular in the heavy-tailed case, moment generating functions are not a means to characterize the distribution, and Theorem 2.2.3 cannot be used to determine whether two random variables have the same distribution.

Let F be a distribution function of a random variable ξ. We denote by $\overline{F}$ the tail function of ξ, that is

$$\overline{F}(x) = 1 - F(x) = \mathbb{P}(\xi > x),$$

for all x.

Definition 3.0.2. If $\overline{F}(x) = \mathbb{P}(\xi > x) > 0$ for all x, we say that F is right-unbounded or that F has a *right-unbounded support.*

Remark. In the case of distributions on $\mathbb{R}^+$ it is common to say that the distribution has unbounded support. In the case of distributions on the whole real line we say that the distribution has a right-unbounded support. These mean the same thing as far as the right tail is concerned.

3.1 Long-tailed distributions

Definition 3.1.1. A distribution F on $\mathbb{R}$ with right-unbounded support is called *long-tailed*, if for any fixed $y > 0$,

$$\frac{\overline{F}(x+y)}{\overline{F}(x)} \to 1, \tag{3.1}$$

as $x \to \infty$.

We write $\overline{F}(x+y) \sim \overline{F}(x)$ denoting the property (3.1). The concept of long-tailedness could be understood for example in the following manner: if the random variable is bigger than some x which is a very large number, it is approximately as likely to exceed an even bigger number, that is $x + y$.

3.2 Subexponential distributions

The references for this section are also [Embrechts et al., 2008, Chapter 1] and [Foss et al., 2011, Chapters 1 and 3].

All heavy-tailed distributions likely to be encountered in practice are not just long-tailed but hold even stronger regularity property, that is subexponentiality. What this means is that the tail behaviour is good under the operation of convolution, see [Foss et al., 2011, Chapter 3]. Now we move on to defining what it means that a distribution is subexponential.

Definition 3.2.1. A distribution F on $\mathbb{R}^+$ with unbounded support is called *subexponential*, if

$$\overline{F^{*2}}(x) = \overline{F * F}(x) \sim 2\overline{F}(x), \tag{3.2}$$

as $x \to \infty$.

For subexponential distributions, Condition (3.2) actually holds for every $n \geq 2$. In practice though it is far more convenient to have the definition formulated for one value only, in the situation where checking the definition is needed.

We denote the class of subexponential distributions by $\mathcal{S}$. Subexponentiality is a tail property in the class of distributions on $\mathbb{R}^+$, that is, it only depends on the (right) tail of the distribution. It is a well-recognised fact that any subexponential distribution is heavy-tailed and moreover, long-tailed.

For a long-tailed distribution it holds that $\overline{F}(x)e^{\lambda x} \to \infty$ for all $\lambda > 0$, as $x \to \infty$. It follows from this that the tail function of a long-tailed distribution $\overline{F}$ decays to zero slower than the tail of any exponential distribution function, that is $e^{-\lambda x}$, $\lambda > 0$. This is especially true for subexponential distributions. This observation originally motivated the term subexponential. As pointed out in [Foss et al., 2011, Section 3.1] though, the term is nowadays used in a more restrictive sense, as in Definition 3.2.1.

Subexponentiality of F is equivalent to the demand that, for fixed n,

$$\mathbb{P}(\max(\xi_1, \ldots, \xi_n) > x) \sim \mathbb{P}(\xi_1 + \cdots + \xi_n > x), \tag{3.3}$$

as $x \to \infty$. This can be interpreted in the following way: the only significant situation where $S_n = \xi_1 + \cdots + \xi_n$ exceeds x (a large number) is when one of the summands itself exceeds x. This is known as the principle of a single big jump or catastrophe principle. So what is characteristic of this class of distributions $\mathcal{S}$ is that a single increment might dominate the magnitude of the sum.

Even though many heavy-tailed distributions of practical applications are subexponential, [Embrechts et al., 2008, Section 1.4] gives an example of a heavy-tailed distribution, which is not subexponential. Such distribution is the so-called Peter and Paul distribution. Let us explain the situation. Consider a game where Peter tosses a fair coin until it falls head for the first time. Suppose this happens at the kth toss. Then Peter receives 2^k euros from Paul. The distribution function F of Peter's gain is

$$F(x) = \sum_{k: 2^k \le x} 2^{-k},$$

for $x \ge 0$. Since for all natural numbers k,

$$\frac{\overline{F}(2^k - 1)}{\overline{F}(2^k)} = 2,$$

it follows that F is not long-tailed nor subexponential.

Remark. In [Embrechts et al., 2008], distributions are defined on sets different from ours.

Definition 3.2.1 was formulated merely for the distributions on the non-negative half-line $\mathbb{R}^+$. We would like to extend the concept of subexponential distributions to the whole real line. Before defining the whole-line subexponentiality, we motivate the definition with the following lemma.

Lemma 3.2.2. *Let F be a distribution on $\mathbb{R}$ and let ξ be a random variable with distribution F. Then the following are equivalent:*

(i) F is long-tailed and $\overline{F * F}(x) \sim 2\overline{F}(x)$, as $x \to \infty$.

(ii) The distribution F^+ of $\xi^+ = \max(\xi, 0)$ is subexponential.

(iii) The conditional distribution $G(B) = \mathbb{P}(\xi \in B \mid \xi \geq 0)$ is subexponential.

Proof. We refer the reader to [Foss et al., 2011, Section 3.2] for the proof. □

Now we can define subexponentiality on $\mathbb{R}$, following the representation of [Foss et al., 2011, Section 3.2].

Definition 3.2.3. Let F be a distribution on $\mathbb{R}$ with right-unbounded support. We say that F is *whole-line subexponential*, and write $F \in \mathcal{S}_{\mathbb{R}}$, if F is long-tailed and

$$\overline{F * F}(x) \sim 2\overline{F}(x),$$

as $x \to \infty$. Equivalently, a random variable ξ has a whole-line subexponential distribution if ξ^+ has a subexponential distribution.

In [Foss et al., 2011, Section 3.2] it is stated that $\mathcal{S} \subseteq \mathcal{S}_{\mathbb{R}}$ and furthermore, $\mathcal{S}_{\mathbb{R}}$ is a subset of the class of long-tailed distributions.

3.3 Strong subexponential distributions

The reference for this section is also [Foss et al., 2011, Chapter 3].

In some applications, those concerning the behaviour of the maxima of a random walk with heavy-tailed increments for example, we require a little more regularity when it comes to their tails. Sufficient regularity is satisfied by those distributions that are of the form of Definition 3.3.1, see [Foss et al., 2011, Section 3.4].

Definition 3.3.1. A distribution F on $\mathbb{R}$ with right-unbounded support and finite expectation belongs to the *class* $\mathcal{S}^*$ if

$$\int_0^x \overline{F}(x-y)\overline{F}(y)\mathrm{d}y \sim 2a\overline{F}(x),$$

as $x \to \infty$, where $a = \int_0^\infty \overline{F}(y)\mathrm{d}y$ is the expectation of $F^+ = F\mathbb{1}_{\mathbb{R}^+}$.

It is known that any distribution from class $\mathcal{S}^*$ is (whole-line) subexponential. For a probabilistic proof on the matter, see Corollary 1 from [Foss and Zachary, 2003].

Remark. When we talk about distributions on $\mathbb{R}$ and it is clear from the context, we might just say that the distribution is subexponential, even though whole-line subexponentiality is meant.

The two classes, the class of subexponential distributions $\mathcal{S}$ and the subclass of subexponential distributions $\mathcal{S}^*$, are found to be relatively similar. However, there exists subexponential distributions that are not in $\mathcal{S}^*$. Most likely motivated by this, for example in [Foss et al., 2011], the class $\mathcal{S}^*$ is called *strong subexponential.* In particular, the class $\mathcal{S}^*$ contains all the heavy-tailed distributions likely to be encountered in practice. Some well-known distributions of the class $\mathcal{S}^*$ are for instance Pareto distribution along with log-normal distribution.

As the Pareto distribution is strong subexponential, it implies it is also heavy-tailed. We shall demonstrate as to why the moment generating function does not exist as a finite number for strictly positive values of the argument. Hence, let us examine the expectation $\mathbb{E}(e^{s\xi})$, where $s > 0$. Let $\alpha > 0$ be the shape parameter and $r > 0$ the scale parameter. Let F be the distribution function of a Pareto(α, r) distributed random variable ξ, that is

$$F_\xi(x) = \begin{cases} 1 - \left(\frac{r}{x}\right)^\alpha, & x \geq r \\ 0, & \text{otherwise.} \end{cases}$$

For the density function we have

$$f_\xi(x) = \begin{cases} \frac{\alpha r^\alpha}{x^{\alpha+1}}, & x \geq r \\ 0, & \text{otherwise.} \end{cases}$$

We then write

$$M_\xi(s) = \mathbb{E}(e^{s\xi}) = \int_{-\infty}^{\infty} e^{sx} \mathrm{d}F_\xi(x) = \int_{-\infty}^{\infty} e^{sx} f_\xi(x) \mathrm{d}x = \int_{r}^{\infty} e^{sx} \frac{\alpha r^\alpha}{x^{\alpha+1}} \mathrm{d}x. \tag{3.4}$$

The integral in (3.4) might converge if the integrand converges. Let us look at the limit

$$\lim_{x\to\infty} \frac{e^{sx}}{x^{\alpha+1}} \tag{3.5}$$

more closely. After multiple applications of L'Hôpital's rule it is seen that the limit in (3.5) equals ∞. Since the integrand diverges as x tends to infinity, the whole integral must diverge as well. Thus the moment generating function $M_\xi(s)$ of Pareto distribution does not exist as a finite number for any $s > 0$.

How about for $s \leq 0$? For such values of s, the limit in (3.5) equals 0 and the integral in (3.4) converges. When $s = 0$, (3.4) becomes

$$\alpha r^\alpha \int_r^\infty \frac{1}{x^{\alpha+1}} \mathrm{d}x.$$

We continue writing this as

$$
\begin{aligned}
\alpha r^{\alpha} \int_{r}^{\infty} \frac{1}{x^{\alpha+1}} \mathrm{d}x &= \alpha r^{\alpha} \int_{r}^{\infty} x^{-\alpha-1} \mathrm{d}x \\
&= \alpha r^{\alpha} \Big/_{r}^{\infty} \frac{1}{-\alpha} x^{-\alpha} \\
&= \alpha r^{\alpha} (0 - \frac{1}{-\alpha} r^{-\alpha}) \\
&= \alpha r^{\alpha} \frac{1}{\alpha r^{\alpha}} \\
&= 1.
\end{aligned}
$$

When $s < 0$, (3.4) becomes

$$
\begin{aligned}
\alpha r^{\alpha} \int_{r}^{\infty} e^{sx} x^{-\alpha-1} \mathrm{d}x &= \alpha r^{\alpha} \int_{-rs}^{\infty} \left| \frac{1}{-s} \right| e^{-x} \left(\frac{1}{-s} x \right)^{-\alpha-1} \mathrm{d}x \\
&= \alpha r^{\alpha} \left(\frac{1}{-s} \right)^{-\alpha} \int_{-rs}^{\infty} e^{-x} x^{-\alpha-1} \mathrm{d}x \\
&= \alpha r^{\alpha} (-s)^{\alpha} \Gamma(-\alpha, -rs) \\
&= \alpha (-rs)^{\alpha} \Gamma(-\alpha, -rs),
\end{aligned}
$$

where by Γ we denoted the upper incomplete gamma function.

So for all values of the argument s, we have

$$
M_{\xi}(s) = \begin{cases} 1, & s = 0 \\ \alpha(-rs)^{\alpha} \Gamma(-\alpha, -rs), & s < 0 \\ \text{does not exist,} & \text{otherwise.} \end{cases}
$$

3.4 Technical auxiliary results for asymptotics

This section focuses on preparation for the upcoming asymptotic results. We will present results needed for the proof of Theorem 5.2.1.

Next we will present a lemma that shows that asymptotic equivalence for (at infinity) strictly positive functions is an equivalence relation.

Lemma 3.4.1. *The asymptotic equivalence of (at infinity) strictly positive functions is an equivalence relation.*

Proof. Suppose that $u(x), (x)$ and $w(x)$ are functions from $\mathbb{R}$ to $(0,\infty)$. We shall first prove the reflexive property. As $\frac{u(x)}{u(x)} = 1$ for all $x \in \mathbb{R}$, in particular it holds that

$$\lim_{x\to\infty} \frac{u(x)}{u(x)} = 1 \Rightarrow u(x) \sim u(x).$$

We will proceed with the symmetric property. Suppose that $u(x) \sim v(x)$. We aim to show that $v(x) \sim u(x)$.

$$\lim_{x\to\infty} \frac{v(x)}{u(x)} = 1 \cdot \lim_{x\to\infty} \frac{v(x)}{u(x)} = \lim_{x\to\infty} \frac{u(x)}{v(x)} \lim_{x\to\infty} \frac{v(x)}{u(x)} = \lim_{x\to\infty} \frac{u(x)}{v(x)} \frac{v(x)}{u(x)} = 1.$$

That is, $v(x) \sim u(x)$.

Lastly, we will prove the transitive property. Suppose now that $u(x) \sim v(x)$ and $v(x) \sim w(x)$. We wish to prove that $u(x) \sim w(x)$.

$$\lim_{x\to\infty} \frac{u(x)}{w(x)} = \lim_{x\to\infty} \frac{u(x)v(x)}{w(x)v(x)} = \lim_{x\to\infty} \frac{u(x)}{v(x)} \lim_{x\to\infty} \frac{v(x)}{w(x)} = 1 \cdot 1 = 1.$$

Therefore $u(x) \sim w(x)$. $\square$

The next theorem is mentioned in [Denisov et al., 2010, Section 2].

Theorem 3.4.2. *Let $\xi, \xi_1, \xi_2, \ldots$ be independent random variables with common distribution F. Assume that $F \in \mathcal{S}^*$ and $\mathbb{E}(\xi) < 0$. Write S_n for the sum $\xi_1 + \cdots + \xi_n$. Denote by M_n the maximum $\max_{0\le i\le n} S_i$. Then, as $x \to \infty$ and uniformly in $n \ge 1$,*

$$\mathbb{P}(M_n > x) \sim \frac{1}{|\mathbb{E}(\xi)|} \int_x^{x+n|\mathbb{E}(\xi)|} \overline{F}(y)\mathrm{d}y.$$

Theorem 3.4.3. *(Continuous version of Fatou's lemma.) Assume that $\{f_x\}_{x\ge 0}$ is a collection of non-negative μ-measurable functions and $\liminf_{x\to\infty} f_x$ is μ-measurable. Then it holds that*

$$\int \liminf_{x\to\infty} f_x \mathrm{d}\mu \le \liminf_{x\to\infty} \int f_x \mathrm{d}\mu.$$

Proof. In what follows, assume that $n \in \mathbb{N}^+$ and $x, x_n \in \mathbb{R}^+$. To begin the proof, we make a claim: there exists a sequence $\{x_n\}$ of non-negative real numbers such that $x_n \to \infty$, as $n \to \infty$, $x_n \ge n$ and

$$\lim_{n\to\infty} \int f_{x_n} \mathrm{d}\mu = \liminf_{x\to\infty} \int f_x \mathrm{d}\mu. \tag{3.6}$$

First, write $\int f_x \mathrm{d}\mu = I_x$. Now equation (3.6) becomes $\lim_{n\to\infty} I_{x_n} = \liminf_{x\to\infty} I_x$. Denote

$$b = \liminf_{x\to\infty} I_x = \lim_{n\to\infty} \inf_{x\geq n} I_x.$$

Note, that $\{\inf_{x\geq n} I_x\}$ is an increasing sequence with respect to n.

i) To prove the claim we made, assume first that $\liminf_{x\to\infty} I_x = b < \infty$. We aim to show that the previously described sequence $\{x_n\}$ exists, such that it also holds that $\lim_{n\to\infty} I_{x_n} = b$. Then (3.6) holds in the case where both the left- and right-hand side are finite.

Fix an arbitrary $n \in \mathbb{N}^+$. Consider the set $\{I_x \mid x \geq n\}$. There exists x_n such that $I_{x_n} \in \{I_x \mid x \geq n\}$ and additionally

$$\inf_{x\geq n} I_x \leq I_{x_n} \leq \inf_{x\geq n} I_x + \frac{1}{n}. \tag{3.7}$$

Note that since $I_{x_n} \in \{I_n, I_{n+1}, I_{n+2}, ...\}$, it means that $x_n \geq n$. Since $n \to \infty$, it follows that $x_n \to \infty$, as $n \to \infty$.

The inequalities of (3.7) are the result of the definition of infimum. Write $\inf_{x\geq n} I_x = c$ for the infimum. When we move up from c a positive amount ($\frac{1}{n}$ here), there must be an element that is in the interval $[c, c+\frac{1}{n})$. If there were not, c would not be the infimum, instead some larger number would. But as c is the infimum, we can find a number I_{x_n} that is greater than or equal to c and less than or equal to $c + \frac{1}{n} = \inf_{x\geq n} I_x + \frac{1}{n}$. This does not hold for all possible values of x_n, but we were able to choose the value of x_n to be such that it holds.

We assumed that $\lim_{n\to\infty} \inf_{x\geq n} I_x = b$ and therefore also $\lim_{n\to\infty} (\inf_{x\geq n} I_x + \frac{1}{n}) = b$. By applying the squeeze theorem to (3.7), $\lim_{n\to\infty} I_{x_n} = b$, which yields the result.

ii) Assume then that $b = \infty$. For all $n \in \mathbb{N}^+$ there exists $I_{x_n} \in \{I_x : x \geq n\}$ such that $\inf_{x\geq n} I_x \leq I_{x_n}$. Also here it holds that $x_n \geq n$. Therefore

$$\infty = b = \lim_{n\to\infty} \inf_{x\geq n} I_x \leq \lim_{n\to\infty} I_{x_n},$$

which implies that $\lim_{n\to\infty} I_{x_n}$ itself equals infinity. Thus $\lim_{n\to\infty} I_{x_n} = \liminf_{x\to\infty} I_x$ also in the case where $b = \infty$.

We have now formulated a sequence $\{x_n\}$ of non-negative real numbers such that $x_n \to \infty$, as $n \to \infty$, $x_n \geq n$ and (3.6) holds. Now that we have proved the claim we made in the beginning, we proceed with the proof of the theorem.

We have

$$\liminf_{x\to\infty} f_x = \lim_{k\to\infty} \inf_{x\geq k} f_x \leq \lim_{k\to\infty} \inf_{x_n\geq k} f_{x_n} \leq \lim_{k\to\infty} \inf_{n\geq k} f_{x_n} = \liminf_{n\to\infty} f_{x_n}, \tag{3.8}$$

where $k \in \mathbb{N}$. Hence we may deduce that

$$\int \liminf_{x\to\infty} f_x \mathrm{d}\mu \leq \int \liminf_{n\to\infty} f_{x_n} \mathrm{d}\mu \leq \liminf_{n\to\infty} \int f_{x_n} \mathrm{d}\mu = \lim_{n\to\infty} \int f_{x_n} \mathrm{d}\mu = \liminf_{x\to\infty} \int f_x \mathrm{d}\mu.$$

In the first inequality we applied (3.8) and Fatou's lemma yields the second inequality. The first equality follows when we note that the limit exists and the second equality is (3.6). □

We will need the previous theorem in the proof of the next one.

Theorem 3.4.4. *(Reverse version of the continuous Fatou's lemma) Assume that* $\{f_x\}_{x\geq 0}$ *is a collection of non-negative* μ*-measurable functions. If there exists a non-negative integrable function* g *such that* $f_x \leq g$ *for all* x *and that* $\liminf_{x\to\infty}(g - f_x)$ *is* μ*-measurable, then*

$$\limsup_{x\to\infty} \int f_x \mathrm{d}\mu \leq \int \limsup_{x\to\infty} f_x \mathrm{d}\mu.$$

Proof. Assume $x \in \mathbb{R}^+$ and $n \in \mathbb{N}$. Since $g - f_x$ is non-negative, we get by using Theorem 3.4.3 that

$$\int \liminf_{x\to\infty}(g - f_x)\mathrm{d}\mu \leq \liminf_{x\to\infty} \int g - f_x \mathrm{d}\mu = \liminf_{x\to\infty} \left(\int g \mathrm{d}\mu - \int f_x \mathrm{d}\mu \right). \tag{3.9}$$

Note that

$$\begin{aligned}
\liminf_{x\to\infty}(g + (-f_x)) &= \lim_{n\to\infty} (\inf_{x\geq n} (g + (-f_x))) \\
&= \lim_{n\to\infty} (g + \inf_{x\geq n}(-f_x)) \\
&= \lim_{n\to\infty} (g - \sup_{x\geq n} f_x) \\
&= g - \lim_{n\to\infty} (\sup_{x\geq n} f_x) \\
&= g - \limsup_{x\to\infty} f_x.
\end{aligned}$$

Similarly we deduce that

$$\liminf_{x\to\infty} \left(- \int f_x \mathrm{d}\mu \right) = - \limsup_{x\to\infty} \int f_x \mathrm{d}\mu.$$

Based on these calculations we can write (3.9) as

$$\int g - \limsup_{x\to\infty} f_x \mathrm{d}\mu \leq \int g\mathrm{d}\mu - \limsup_{x\to\infty} \int f_x \mathrm{d}\mu.$$

From this we deduce that

$$-\int \limsup_{x\to\infty} f_x \mathrm{d}\mu \leq -\limsup_{x\to\infty} \int f_x \mathrm{d}\mu \;\Rightarrow\; \int \limsup_{x\to\infty} f_x \mathrm{d}\mu \geq \limsup_{x\to\infty} \int f_x \mathrm{d}\mu,$$

which is the desired result. $\square$

In Chapter 5 we will state asymptotic results and give some proofs of the most important ones, and we shall make use of the results of this section.

Chapter 4

Average behaviour of compound variables

In this chapter we derive expressions for the expectation of the compound variable as well as for the sum variable, where the counting random variable τ does not have to be independent of all ξ_i's, $i = 1, 2, \ldots$. These results form two versions of Wald's identity. The more general version permits us to apply the expectation result to situations where it is only required that the event $\{\tau \leq n\}$ can not depend on ξ_{n+1}.

4.1 Expectation and variance of a compound variable

For the expectation of a general compound variable, we derive the formula in the following lemma. Note that $\mathbb{E}(S_0) = 0$ clearly, when we define $S_0 = 0$.

Lemma 4.1.1. *Assume S_τ is a compound variable and that $\mathbb{E}(\xi) < \infty$ and $\mathbb{E}(\tau) < \infty$. Then*

$$\mathbb{E}(S_\tau) = \mathbb{E}(\tau)\mathbb{E}(\xi).$$

Proof. Since S_τ is a compound variable, we know that the counting variable τ and $\xi_1, \xi_2, \xi_3, \ldots$ are independent. Furthermore, the ξ_i's, $i = 1, 2, 3, \ldots$ are identically distributed random variables.

Under the assumptions, we make the following calculation:

$$\mathbb{E}(S_\tau) = \sum_{t=1}^{\infty} \mathbb{P}(\tau = t)\mathbb{E}(S_\tau \mid \tau = t) = \sum_{t=1}^{\infty} \mathbb{P}(\tau = t)\mathbb{E}(S_t \mid \tau = t).$$

We continue writing this as

$$\begin{aligned}\sum_{t=1}^{\infty}\mathbb{P}(\tau=t)\mathbb{E}(S_t\mid\tau=t)&=\sum_{t=1}^{\infty}\mathbb{P}(\tau=t)\mathbb{E}(S_t)\\&=\sum_{t=1}^{\infty}\mathbb{P}(\tau=t)t\mathbb{E}(\xi)\\&=\mathbb{E}(\tau)\mathbb{E}(\xi).\end{aligned}\tag{4.1}$$

□

In the case where the compound variable depicts the total claim amount, the expectation $\mathbb{E}(S_\tau)$ can be thought of as the amount that the insurance company must be prepared to pay to the policy-holders as a result of the claims they made. Lemma 4.1.1 is sometimes called the basic version of Wald's identity. We will state and prove the more general version of the identity shortly.

The variance of a compound variable is obtained from the formula

$$\mathrm{Var}(S_\tau)=\mathbb{E}({S_\tau}^2)-\mathbb{E}(S_\tau)^2,\tag{4.2}$$

by determining $\mathbb{E}({S_\tau}^2)$ and subtracting the second power of (4.1) from it. Based on Chapter 2 we know that

$$\mathbb{E}({S_\tau}^2)=M_{S_\tau}^{(2)}(0).$$

By Theorem 2.3.3,

$$M_{S_\tau}(s)=M_\tau(c_\xi(s))=M_\tau(\ln(M_\xi(s))).$$

Therefore,

$$M'_{S_\tau}(s)=M'_\tau(\ln(M_\xi(s)))\frac{1}{M_\xi(s)}M'_\xi(s)$$

and

$$\begin{aligned}M_{S_\tau}^{(2)}(s)=&-M'_\tau(\ln(M_\xi(s)))\frac{(M'_\xi(s))^2}{(M_\xi(s))^2}\\&+\frac{1}{M_\xi(s)}\left[M_\tau^{(2)}(\ln(M_\xi(s)))\frac{(M'_\xi(s))^2}{M_\xi(s)}+M'_\tau(\ln(M_\xi(s)))M_\xi^{(2)}(s)\right].\end{aligned}\tag{4.3}$$

Evaluated at $s = 0$, (4.3) becomes

$$M_{S_\tau}^{(2)}(0) = -\mathbb{E}(\tau)(\mathbb{E}(\xi))^2 + \mathbb{E}(\tau^2)(\mathbb{E}(\xi))^2 + \mathbb{E}(\tau)\mathbb{E}(\xi^2).$$

To finalize determining (4.2), we make the subtraction

$$\begin{aligned}\operatorname{Var}(S_\tau) &= -\mathbb{E}(\tau)(\mathbb{E}(\xi))^2 + \mathbb{E}(\tau^2)(\mathbb{E}(\xi))^2 + \mathbb{E}(\tau)\mathbb{E}(\xi^2) - (\mathbb{E}(\tau)\mathbb{E}(\xi))^2 \\ &= \left[\mathbb{E}(\tau^2) - \mathbb{E}(\tau) - (\mathbb{E}(\tau))^2\right](\mathbb{E}(\xi))^2 + \mathbb{E}(\tau)\mathbb{E}(\xi^2) \\ &= \mathbb{E}(\tau)\operatorname{Var}(\xi) + (\mathbb{E}(\xi))^2\operatorname{Var}(\tau). \end{aligned} \tag{4.4}$$

4.1.1 Expectation and variance of a compound Poisson variable

When considering a compound Poisson variable S_τ, with parameter (λ, F), we determine the expectation as we would for any other compound variable, taking into account the Poisson parameter as the expectation of τ. Hence we have that

$$\mathbb{E}(S_\tau) = \lambda\mathbb{E}(\xi).$$

Recall that since $\tau \sim \text{Poisson}(\lambda)$,

$$\mathbb{E}(\tau) = \operatorname{Var}(\tau) = \lambda.$$

Thus for the variance of S_τ, a compound Poisson variable with parameter (λ, F), we substitute in (4.4) what we know and deduce that

$$\operatorname{Var}(S_\tau) = \lambda(\mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2) + (\mathbb{E}(\xi))^2\lambda = \lambda\mathbb{E}(\xi^2).$$

4.2 Wald's identity

Lemma 4.2.1. *(Wald's identity.) Let τ be a counting random variable such that for every $n \in \mathbb{N}$,*

$$\textit{the event } \{\tau \le n\} \textit{ does not depend on } \xi_{n+1}. \tag{4.5}$$

Then for $S_\tau = \xi_1 + \cdots + \xi_\tau$, where $\xi_1, \xi_2, \xi_3, \ldots$ are independent, identically distributed random variables, it holds that

$$\mathbb{E}(S_\tau) = \mathbb{E}(\tau)\mathbb{E}(\xi),$$

provided that both $\mathbb{E}(\tau)$ and $\mathbb{E}(|\xi|)$ are finite.

Remark. The result holds also when we assume that τ is a stopping time.

Remark. About the notation, we write

$$\mathbb{E}(\xi_k; \tau = n) = \mathbb{E}(\xi_k \mathbb{1}(\tau = n))$$

for the notation found in [Foss et al., 2011].

Now we can begin the proof of Lemma 4.2.1.

Proof. The proof follows the idea of the one in [Foss et al., 2011]. We will begin by making the following decomposition:

$$\begin{aligned}
\mathbb{E}(S_\tau) &= \sum_{n=1}^{\infty} \mathbb{E}(S_\tau \mathbb{1}(\tau = n)) \\
&= \sum_{n=1}^{\infty} \mathbb{E}(S_\tau \mid \tau = n)\mathbb{P}(\tau = n) \\
&= \mathbb{E}(S_\tau \mid \tau = 1)\mathbb{P}(\tau = 1) + \mathbb{E}(S_\tau \mid \tau = 2)\mathbb{P}(\tau = 2) + \mathbb{E}(S_\tau \mid \tau = 3)\mathbb{P}(\tau = 3) + \cdots \\
&= \mathbb{E}(\xi_1 \mid \tau = 1)\mathbb{P}(\tau = 1) + \mathbb{E}(\xi_1 + \xi_2 \mid \tau = 2)\mathbb{P}(\tau = 2) \\
&\qquad + \mathbb{E}(\xi_1 + \xi_2 + \xi_3 \mid \tau = 3)\mathbb{P}(\tau = 3) + \cdots \\
&= \mathbb{E}(\xi_1 \mid \tau = 1)\mathbb{P}(\tau = 1) + \mathbb{E}(\xi_1 \mid \tau = 2)\mathbb{P}(\tau = 2) + \mathbb{E}(\xi_2 \mid \tau = 2)\mathbb{P}(\tau = 2) \\
&\qquad + \mathbb{E}(\xi_1 \mid \tau = 3)\mathbb{P}(\tau = 3) + \mathbb{E}(\xi_2 \mid \tau = 3)\mathbb{P}(\tau = 3) + \mathbb{E}(\xi_3 \mid \tau = 3)\mathbb{P}(\tau = 3) + \cdots
\end{aligned}$$

Writing this under the sum sign, we have

$$\begin{aligned}
&\sum_{k=1}^{1} \mathbb{E}(\xi_k \mid \tau = 1)\mathbb{P}(\tau = 1) + \sum_{k=1}^{2} \mathbb{E}(\xi_k \mid \tau = 2)\mathbb{P}(\tau = 2) \\
&\qquad + \sum_{k=1}^{3} \mathbb{E}(\xi_k \mid \tau = 3)\mathbb{P}(\tau = 3) + \cdots \\
&= \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathbb{E}(\xi_k \mid \tau = n)\mathbb{P}(\tau = n) \\
&= \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathbb{E}(\xi_k \mathbb{1}(\tau = n)). \qquad (4.6)
\end{aligned}$$

Next we will justify interchanging the order of summation in (4.6). Taking absolute values of the summands we have

$$\sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathbb{E}(|\xi_k| \mathbb{1}(\tau = n)) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mathbb{E}(|\xi_k| \mathbb{1}(\tau = n))\mathbb{1}(k \le n)$$

and now that the summands are non-negative, the application of Tonelli's theorem is justified and we arrive at

$$\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\mathbb{E}(|\xi_k|\mathbb{1}(\tau=n))\mathbb{1}(k\le n)=\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\mathbb{E}(|\xi_k|\mathbb{1}(\tau=n)). \tag{4.7}$$

Writing out the outer sum to see what the formula is like, (4.7) yields

$$\begin{aligned}
&\sum_{n=1}^{\infty}\mathbb{E}(|\xi_1|\mathbb{1}(\tau=n))+\sum_{n=2}^{\infty}\mathbb{E}(|\xi_2|\mathbb{1}(\tau=n))+\sum_{n=3}^{\infty}\mathbb{E}(|\xi_3|\mathbb{1}(\tau=n))+\cdots\\
&=\mathbb{E}(|\xi_1|\mathbb{1}(\tau=1))+\mathbb{E}(|\xi_1|\mathbb{1}(\tau=2))+\mathbb{E}(|\xi_1|\mathbb{1}(\tau=3))+\cdots\\
&\qquad+\mathbb{E}(|\xi_2|\mathbb{1}(\tau=2))+\mathbb{E}(|\xi_2|\mathbb{1}(\tau=3))+\mathbb{E}(|\xi_2|\mathbb{1}(\tau=4))+\cdots\\
&\qquad\qquad+\mathbb{E}(|\xi_3|\mathbb{1}(\tau=3))+\mathbb{E}(|\xi_3|\mathbb{1}(\tau=4))+\mathbb{E}(|\xi_3|\mathbb{1}(\tau=5))+\cdots\\
&\qquad\qquad\qquad\vdots\\
&=\mathbb{E}(|\xi_1|(\mathbb{1}(\tau=1)+\mathbb{1}(\tau=2)+\mathbb{1}(\tau=3)+\cdots))\\
&\qquad+\mathbb{E}(|\xi_2|(\mathbb{1}(\tau=2)+\mathbb{1}(\tau=3)+\mathbb{1}(\tau=4)+\cdots))\\
&\qquad\qquad+\mathbb{E}(|\xi_3|(\mathbb{1}(\tau=3)+\mathbb{1}(\tau=4)+\mathbb{1}(\tau=5)+\cdots))\\
&\qquad\qquad\qquad\vdots\\
&=\mathbb{E}(|\xi_1|\mathbb{1}(\tau\ge 1))+\mathbb{E}(|\xi_2|\mathbb{1}(\tau\ge 2))+\mathbb{E}(|\xi_3|\mathbb{1}(\tau\ge 3))+\cdots\\
&=\sum_{k=1}^{\infty}\mathbb{E}(|\xi_k|\mathbb{1}(\tau\ge k)).
\end{aligned}\tag{4.8}$$

If in addition (4.8) is finite, then the change in summation order is justified also in (4.6).

By Condition (4.5) we mean that the random variable $\mathbb{1}(\tau\le n)$ is independent of the random variable ξ_{n+1}, which implies that

$$\mathbb{E}(\xi_{n+1}\mathbb{1}(\tau\le n))=\mathbb{E}(\xi_{n+1})\mathbb{E}(\mathbb{1}(\tau\le n))=\mathbb{E}(\xi_{n+1})\mathbb{P}(\tau\le n).$$

By Condition (4.5), the event $\{\tau\le k-1\}$ does not depend on ξ_k i.e. the random variables $\mathbb{1}(\tau\le k-1)$ and ξ_k are independent. Furthermore, the random variable $\mathbb{1}(\tau\le k-1)$ does not depend on the measurable transformation of ξ_k, that is $|\xi_k|$. Let us consider $1-\mathbb{1}(\tau\le k-1)=\mathbb{1}(\tau\ge k)$. As a measurable transformation of $\mathbb{1}(\tau\le k-1)$, the random variable $\mathbb{1}(\tau\ge k)$ does not depend on $|\xi_k|$ (or ξ_k) either. We will apply this argument in the first equality of the following calculation, where we check that (4.8) is finite. Let us continue writing (4.8) as follows.

$$
\begin{aligned}
\sum_{k=1}^{\infty} \mathbb{E}(|\xi_k| \mathbb{1}(\tau \geq k)) &= \sum_{k=1}^{\infty} \mathbb{E}(|\xi_k|) \mathbb{P}(\tau \geq k) \\
&= \mathbb{E}(|\xi|) \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} \mathbb{P}(\tau = j) \\
&= \mathbb{E}(|\xi|) \left[\mathbb{P}(\tau = 1) + \mathbb{P}(\tau = 2) + \mathbb{P}(\tau = 3) + \cdots \right] \\
&\qquad + \mathbb{E}(|\xi|) \left[\mathbb{P}(\tau = 2) + \mathbb{P}(\tau = 3) + \mathbb{P}(\tau = 4) + \cdots \right] \\
&\qquad\qquad + \mathbb{E}(|\xi|) \left[\mathbb{P}(\tau = 3) + \mathbb{P}(\tau = 4) + \mathbb{P}(\tau = 5) + \cdots \right] \\
&\qquad\qquad\qquad \vdots \\
&= \mathbb{E}(|\xi|) \left[1 \cdot \mathbb{P}(\tau = 1) + 2 \cdot \mathbb{P}(\tau = 2) + 3 \cdot \mathbb{P}(\tau = 3) + \cdots \right] \\
&= \mathbb{E}(|\xi|) \sum_{k=1}^{\infty} k \mathbb{P}(\tau = k), \\
&= \mathbb{E}(|\xi|) \mathbb{E}(\tau) < \infty. \qquad (4.9)
\end{aligned}
$$

The inequality (4.9) holds because both factors were assumed to be finite. Therefore turns out that (4.8) indeed is finite.

Now it is justified to change the order of summation in (4.6) using Fubini's theorem, and by performing similar calculations as above, we have

$$
\mathbb{E}(S_\tau) = \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathbb{E}(\xi_k \mathbb{1}(\tau = n)) = \sum_{k=1}^{\infty} \mathbb{E}(\xi_k \mathbb{1}(\tau \geq k)).
$$

Using again the independence of the event $\{\tau \geq k\}$ and ξ_k, we obtain the desired result

$$
\mathbb{E}(S_\tau) = \sum_{k=1}^{\infty} \mathbb{E}(\xi_k \mathbb{1}(\tau \geq k) = \sum_{k=1}^{\infty} \mathbb{E}(\xi_k) \mathbb{P}(\tau \geq k) = \mathbb{E}(\tau)\mathbb{E}(\xi).
$$

□

Remark. Note, that Wald's identity does not assume the independence of τ and ξ_i's, $i = 1, 2, 3, \ldots$, as opposed to what we assumed in the beginning of this chapter when determining the expectation of a general compound variable S_τ.

Even though it was assumed in the definition of the compound variable, in Section 2.3 we mentioned that the independence of $\tau, \xi_1, \xi_2, \xi_3, \ldots$ might not be realistic. As a consequence the number of claims variable can affect the size of a single claim. In such situations we use Wald's identity to determine the expectation of the random sum S_τ, provided that Condition (4.5) is fulfilled.

An example, from [Daykin et al., 1994, Chapter 3], of a situation where Wald's identity would be applicable is the case of wind storms. The wind storms cause the number of claims to increase drastically every now and then. In addition to very large claims occurring, the small ones will form the majority as opposed to normal circumstances. Therefore the (large) number of claims variable affects the sizes of the claims, and we have dependence there.

Wald's identity has therefore in a sense a broader field of usage due to weaker assumptions about the independence between the summands and the number of summands. It can be applied to situations with dependence between τ and the increments as well as to compound variables where there is no dependence. In particular, the number of summands of a compound variable satisfies Condition (4.5).

Chapter 5

Asymptotic behaviour of compound variables

We commence this chapter by stating and proving an asymptotic result for a tail of a sum variable, where the number of summands is not a random variable but a deterministic number. Then we will start the treatment of the asymptotic results for the tails of compound variables. Theorems 5.2.1, 5.3.1 as well as 5.3.4 are adapted from [Denisov et al., 2010].

Assume in this chapter that $\xi, \xi_1, \xi_2, \ldots$ are independent, identically distributed random variables. Additionally, assume that their expectation is finite, that is $\mathbb{E}(\xi) < \infty$. Moreover, assume that their common distribution F is right-unbounded. Assume further that F has a heavy (right) tail.

Let $S_0 = 0$ and denote by S_n the sum $\xi_1 + \cdots + \xi_n$, where $n = 1, 2, \ldots$. Furthermore, let $M_n = \max\limits_{0 \le i \le n} S_i$ be the maximum of the sum up to time n. Write F^{*n} for the distribution of S_n. Let τ be a counting random variable along with finite expectation, $\mathbb{E}(\tau) < \infty$.

5.1 First asymptotic result

We shall formulate the first asymptotic result for a sum variable S_n, where the number of summands n is not random.

Remark. The following result is actually the other direction of the claim we made in Section 3.2 that subexponentiality is equivalent to Condition (3.3).

Theorem 5.1.1. *Assume that $\xi_1, \xi_2, \xi_3, \ldots$ have distribution function F that is subexponential. Then the following asymptotic equivalence is true:*

$$\mathbb{P}(S_n > x) \sim \mathbb{P}(M'_n > x), \tag{5.1}$$

as $x \to \infty$, where $M_n' = \max(\xi_1, \ldots, \xi_n)$.

Proof. We have that

$$\mathbb{P}(M_n' > x) = 1 - \mathbb{P}(M_n' \leq x) = 1 - \mathbb{P}(\xi_1 \leq x, \ldots, \xi_n \leq x) \tag{5.2}$$

and by independence $\mathbb{P}(\xi_1 \leq x, \ldots, \xi_n \leq x) = \mathbb{P}(\xi_1 \leq x) \cdots \mathbb{P}(\xi_n \leq x)$. We can continue writing the formula (5.2) in the following manner:

$$\begin{aligned}
\mathbb{P}(M_n' > x) &= 1 - \mathbb{P}(\xi_1 \leq x) \cdots \mathbb{P}(\xi_n \leq x) \\
&= 1 - F^n \\
&= 1 - F + F - F^2 + F^2 - \cdots - F^{n-2} + F^{n-2} - F^{n-1} + F^{n-1} - F^n \\
&= F^0 - FF^0 + F^1 - FF^1 + F^2 - \cdots - FF^{n-3} + F^{n-2} - FF^{n-2} \\
&\qquad + F^{n-1} - FF^{n-1} \\
&= (1-F)F^0 + (1-F)F^1 + \cdots + (1-F)F^{n-2} + (1-F)F^{n-1} \\
&= \overline{F} \sum_{k=0}^{n-1} F^k.
\end{aligned}$$

For all powers of F, we wrote $F^n(x) = F^n$ for the sake of clarity.

Distribution functions are increasing functions with $\lim_{x \to \infty} F(x) = 1$. Therefore

$$\lim_{x \to \infty} \sum_{k=0}^{n-1} F^k(x) = n.$$

Hence we have the following asymptotic equivalence:

$$\overline{F}(x) \sum_{k=0}^{n-1} F^k(x) \sim n\overline{F}(x),$$

as $x \to \infty$. Thus we have deduced that

$$\mathbb{P}(M_n' > x) \sim n\overline{F}(x). \tag{5.3}$$

It is known that $\mathbb{P}(S_n > x) = \overline{F^{*n}}(x)$, and therefore the ratio

$$\frac{\mathbb{P}(S_n > x)}{\mathbb{P}(M_n' > x)} \tag{5.4}$$

becomes

$$\frac{\overline{F^{*n}}(x)}{\mathbb{P}(M_n' > x)},$$

which is by (5.3) asymptotically equivalent to

$$\frac{\overline{F^{*n}}(x)}{n\overline{F}(x)}. \tag{5.5}$$

Since F was from the class of subexponential functions, (5.5) is asymptotically equivalent to 1. By the transitivity property of asymptotic equivalence, also the ratio in (5.4) is asymptotically equivalent to 1. We have therefore shown that (5.1) holds. □

5.2 Asymptotic tail approximation: negative expectation

The following theorem is part i) of Theorem 1 of [Denisov et al., 2010].

Theorem 5.2.1. *Assume that a counting random variable τ is independent of $\{\xi_n\}$. Let $F \in \mathcal{S}^*$. If $\mathbb{E}(\xi) < 0$, then*

$$\mathbb{P}(S_\tau > x) \sim \mathbb{P}(M_\tau > x) \sim \mathbb{E}(\tau)\overline{F}(x), \tag{5.6}$$

as $x \to \infty$.

Now that we have made the statement, we would like to motivate this result before starting with the proof. We first note that since the expectation of the summands is negative, it would be rather unlikely that the sum exceeds some very large value x. However, if this were to happen, the most likely explanation for this would be that one of the summand was disproportionately large. The occurrence of very large claims is actually relatively common in heavy-tailed distributions.

To have some intuition to back this theorem up, we notice that if there was one extraordinarily large claim amongst the n claims that form S_n, we would have n options to choose from which was the largest claim that in itself exceeded x. The probability of one claim exceeding x is $\overline{F}(x)$. Multiplying these quantities yields $n\overline{F}(x)$. When the number of summands τ is random, not a deterministic number n, the corresponding formula becomes $\mathbb{E}(\tau)\overline{F}(x)$.

We are now in the position to prove Theorem 5.2.1.

Proof of Theorem 5.2.1. The proof loosely follows the one found in [Denisov et al., 2010]. Using the law of total probability and the assumption of τ being independent of $\{\xi_n\}$, we can make the following decomposition:

$$\mathbb{P}(S_\tau > x) = \mathbb{P}(\xi_1 + \cdots + \xi_\tau > x) = \sum_{n=0}^{\infty} \mathbb{P}(\tau = n)\overline{F^{*n}}(x). \tag{5.7}$$

In order to show (5.6), by symmetry and transitivity properties of asymptotic equivalence, it is enough to show

$$\mathbb{P}(S_\tau > x) \sim \mathbb{E}(\tau)\overline{F}(x) \quad \text{and} \quad \mathbb{P}(M_\tau > x) \sim \mathbb{E}(\tau)\overline{F}(x).$$

To do this, we present the proofs for

$$\lim_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} = \mathbb{E}(\tau) \tag{5.8}$$

and

$$\lim_{x\to\infty} \frac{\mathbb{P}(M_\tau > x)}{\overline{F}(x)} = \mathbb{E}(\tau). \tag{5.9}$$

To prove (5.8), we need to show that

(i) $\liminf\limits_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} \geq \mathbb{E}(\tau)$,

(ii) $\limsup\limits_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} \leq \mathbb{E}(\tau)$.

We then have

$$\mathbb{E}(\tau) \leq \liminf_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} \leq \limsup_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} \leq \mathbb{E}(\tau),$$

which implies that $\liminf\limits_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} = \limsup\limits_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} = \lim\limits_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)}$. From this we deduce that $\lim\limits_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} = \mathbb{E}(\tau)$. With a similar technique we obtain (5.9). We notice that instead of showing the four inequalities, it is enough to show two of them. The other two are immediate consequences. We shall first show that

$$\liminf_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} \geq \mathbb{E}(\tau), \tag{5.10}$$

and as $M_\tau \geq S_\tau$, it holds that

$$\liminf_{x\to\infty} \frac{\mathbb{P}(M_\tau > x)}{\overline{F}(x)} \geq \liminf_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)}.$$

Hence it follows that also

$$\liminf_{x\to\infty} \frac{\mathbb{P}(M_\tau > x)}{\overline{F}(x)} \geq \mathbb{E}(\tau).$$

The other inequality we prove is

$$\limsup_{x\to\infty} \frac{\mathbb{P}(M_\tau > x)}{\overline{F}(x)} \leq \mathbb{E}(\tau), \tag{5.11}$$

as it immediately follows that

$$\limsup_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} \leq \mathbb{E}(\tau),$$

due to the fact that $S_\tau \leq M_\tau$.

With proving these two inequalities and deducing their immediate consequences we have shown that (i) and (ii) hold for both the sum S_τ as well as the maximum M_τ, and therefore (5.8) and (5.9) have been proven to be true. That concludes the proof of the theorem.

To show that the inequality (5.10) holds, we first rewrite the left-hand side.

$$\begin{aligned}\liminf_{x\to\infty} \frac{\mathbb{P}(S_\tau > x)}{\overline{F}(x)} &= \liminf_{x\to\infty} \frac{\sum_{n=0}^{\infty} \mathbb{P}(\tau = n)\overline{F^{*n}}(x)}{\overline{F}(x)} \\ &= \liminf_{x\to\infty} \sum_{n=0}^{\infty} \frac{\mathbb{P}(\tau = n)\overline{F^{*n}}(x)}{\overline{F}(x)}.\end{aligned}$$

In order to use the continuous version of Fatou's lemma in the next step, we justify its application. The use of Fatou's lemma is justified when we notice that clearly $f_x(n) = \frac{\mathbb{P}(\tau=n)\overline{F^{*n}}(x)}{\overline{F}(x)}$, $n \in \mathbb{N}$, are non-negative and measurable based on the fact that the measure μ of the lemma is the counting measure in our case. All functions from $\mathbb{N}$ to $\mathbb{R}$ are measurable with respect to the counting measure. The measurability of $g(n) = \liminf_{x\to\infty} \frac{\mathbb{P}(\tau=n)\overline{F^{*n}}(x)}{\overline{F}(x)}$ is guaranteed as well. Also note that we have assumed that $F \in \mathcal{S}^*$ and $\mathcal{S}^* \subset \mathcal{S}_{\mathbb{R}}$ (as stated in [Foss et al., 2011, Section 3.4]) and therefore F is subexponential. That implies in particular that the limit $\lim_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)}$ exists, and furthermore, equals n. We will use this information in the following calculations. We proceed with the application of Fatou's lemma.

$$
\begin{aligned}
\liminf_{x\to\infty} \sum_{n=0}^{\infty} \frac{\mathbb{P}(\tau = n)\overline{F^{*n}}(x)}{\overline{F}(x)} &\geq \sum_{n=0}^{\infty} \liminf_{x\to\infty} \frac{\mathbb{P}(\tau = n)\overline{F^{*n}}(x)}{\overline{F}(x)} \\
&= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n) \liminf_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \\
&= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n) \lim_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \\
&= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n) n \\
&= \mathbb{E}(\tau).
\end{aligned}
$$

We have thus shown (5.10).

Remark. Note, that there are so far no conditions on the sign of the expected value of ξ.

Suppose that $\mathbb{E}(\xi) < 0$. We proceed with showing that the inequality (5.11) holds. We begin by deriving some useful inequalities.

The inequality in (5.12) is obtained upon noticing that $\overline{F}$ is a decreasing function.

$$
\begin{aligned}
\frac{1}{|\mathbb{E}(\xi)|} \int_x^{x+n|\mathbb{E}(\xi)|} \overline{F}(y)\mathrm{d}y &\leq \frac{1}{|\mathbb{E}(\xi)|} \int_x^{x+n|\mathbb{E}(\xi)|} \overline{F}(x)\mathrm{d}y \\
&= \frac{1}{|\mathbb{E}(\xi)|} n|\mathbb{E}(\xi)|\overline{F}(x) \\
&= n\overline{F}(x).
\end{aligned}
\tag{5.12}
$$

By Theorem 3.4.2, uniformly in $n \geq 1$,

$$
\mathbb{P}(M_n > x) \sim \frac{1}{|\mathbb{E}(\xi)|} \int_x^{x+n|\mathbb{E}(\xi)|} \overline{F}(y)\mathrm{d}y \;\Leftrightarrow\; \lim_{x\to\infty} \frac{\mathbb{P}(M_n > x)}{\frac{1}{|\mathbb{E}(\xi)|} \int_x^{x+n|\mathbb{E}(\xi)|} \overline{F}(y)\mathrm{d}y} = 1.
$$

Suppose then that $\varepsilon > 0$. Then there exists x_ε such that when $x > x_\varepsilon$, the following holds:

$$
\left| \frac{\mathbb{P}(M_n > x)}{\frac{1}{|\mathbb{E}(\xi)|} \int_x^{x+n|\mathbb{E}(\xi)|} \overline{F}(y)\mathrm{d}y} - 1 \right| < \varepsilon \;\Rightarrow\; 1 - \varepsilon < \frac{\mathbb{P}(M_n > x)}{\frac{1}{|\mathbb{E}(\xi)|} \int_x^{x+n|\mathbb{E}(\xi)|} \overline{F}(y)\mathrm{d}y} < 1 + \varepsilon.
$$

In particular, the rightmost inequality implies that

$$\mathbb{P}(M_n > x) < (1+\varepsilon)\frac{1}{|\mathbb{E}(\xi)|}\int_x^{x+n|\mathbb{E}(\xi)|}\overline{F}(y)\mathrm{d}y$$

for large enough x. Now we can evaluate

$$\frac{\mathbb{P}(M_n > x)}{\overline{F}(x)} < \frac{(1+\varepsilon)\frac{1}{|\mathbb{E}(\xi)|}\int_x^{x+n|\mathbb{E}(\xi)|}\overline{F}(y)\mathrm{d}y}{\overline{F}(x)} \leq \frac{(1+\varepsilon)n\overline{F}(x)}{\overline{F}(x)} \leq (1+\varepsilon)n. \tag{5.13}$$

This estimate holds for any $\varepsilon > 0$ and any n. The estimate given in (5.13) also holds for all $x > x_\varepsilon$. This implies that

$$\sup_{x\geq k}\frac{\mathbb{P}(M_n > x)}{\overline{F}(x)} \leq (1+\varepsilon)n,$$

for $k > x_\varepsilon$, $k \in \mathbb{N}$. This further implies that

$$\lim_{k\to\infty}\left(\sup_{x\geq k}\frac{\mathbb{P}(M_n > x)}{\overline{F}(x)}\right) \leq (1+\varepsilon)n.$$

Fix n. Since the above inequality holds for any ε, we deduce that

$$\limsup_{x\to\infty}\frac{\mathbb{P}(M_n > x)}{\overline{F}(x)} \leq n.$$

Now we start deriving the inequality (5.11). We write

$$\limsup_{x\to\infty}\frac{\mathbb{P}(M_\tau > x)}{\overline{F}(x)} = \limsup_{x\to\infty}\frac{\sum_{n=0}^{\infty}\mathbb{P}(\tau = n)\mathbb{P}(M_n > x)}{\overline{F}(x)}. \tag{5.14}$$

For M_τ here we made a decomposition similar to (5.7). We would now like to apply the reverse version of the continuous Fatou's lemma. To justify the application, we need to check that

(i) $\frac{\mathbb{P}(\tau=n)\mathbb{P}(M_n>x)}{\overline{F}(x)}$ are non-negative and measurable.
This is true since clearly they are non-negative and all functions are measurable with respect to the counting measure, which in our situation corresponds to μ in the lemma.

(ii) We need to find an integrable function $g \geq 0$ such that

$$\frac{\mathbb{P}(\tau = n)\mathbb{P}(M_n > x)}{\overline{F}(x)} \leq g, \tag{5.15}$$

for all x and

$$\liminf_{x\to\infty} \left(g - \frac{\mathbb{P}(\tau = n)\mathbb{P}(M_n > x)}{\overline{F}(x)} \right)$$

is measurable.

The left-hand side of (5.15) is bounded from above by

$$\mathbb{P}(\tau = n)(1 + \varepsilon)n, \tag{5.16}$$

due to the inequality obtained in (5.13). This upper bound holds for all $x > x_\varepsilon$ and any $\varepsilon > 0$. For $x \leq x_\varepsilon$, the upper bound can be chosen to be

$$\frac{\mathbb{P}(\tau = n)}{\overline{F}(x_\varepsilon)}. \tag{5.17}$$

The function g in the lemma is chosen to be the maximum of these two upper bounds. It is easily seen to be non-negative. The chosen upper bound does not depend on x and it holds for all x. The integrability of g follows from the fact that it is the maximum of two integrable functions: the upper bound in (5.16) is integrable as

$$\sum_{n=0}^{\infty} \mathbb{P}(\tau = n)(1 + \varepsilon)n = (1 + \varepsilon)\mathbb{E}(\tau),$$

which is a constant multiplied with the expectation of τ, which was assumed to be finite. The upper bound (5.17) is also integrable, since

$$\sum_{n=0}^{\infty} \frac{\mathbb{P}(\tau = n)}{\overline{F}(x_\varepsilon)} = \frac{1}{\overline{F}(x_\varepsilon)} \sum_{n=0}^{\infty} \mathbb{P}(\tau = n) < \frac{1}{\overline{F}(x_\varepsilon)} \mathbb{E}(\tau).$$

Also we know that

$$\liminf_{x\to\infty} \max \left(\mathbb{P}(\tau = n)(1 + \varepsilon)n, \tfrac{1}{\overline{F}(x_\varepsilon)} \mathbb{P}(\tau = n) \right) - \frac{\mathbb{P}(\tau = n)\mathbb{P}(M_n > x)}{\overline{F}(x)}$$

(corresponding to $\liminf_{x\to\infty} g - f_x$ of the lemma) is measurable with respect to the counting measure.

We have now justified the use of reverse continuous Fatou, and (5.14) becomes

$$\begin{aligned}
\limsup_{x\to\infty} \sum_{n=0}^{\infty} \frac{\mathbb{P}(\tau = n)\mathbb{P}(M_n > x)}{\overline{F}(x)} &\le \sum_{n=0}^{\infty} \limsup_{x\to\infty} \frac{\mathbb{P}(\tau = n)\mathbb{P}(M_n > x)}{\overline{F}(x)} \\
&= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n) \limsup_{x\to\infty} \frac{\mathbb{P}(M_n > x)}{\overline{F}(x)} \\
&\le \sum_{n=0}^{\infty} \mathbb{P}(\tau = n)n \\
&= \mathbb{E}(\tau),
\end{aligned}$$

and we have shown the inequality (5.11) to hold. This concludes the proof. □

5.3 Asymptotic tail approximation: non-negative expectation

Reference for this section is the article [Denisov et al., 2010].

Applying the asymptotics obtained in Theorems 5.3.1 and 5.3.4 would be considered for example in the case where S_τ represents the total claim amount. This is due to the non-negativeness assumption of the increments ξ_i, $i = 1, 2, \dots,$ which we found to be natural in the case of total claim amounts. These following results are also applicable in such situations where the non-negativeness assumption for the increments is not satisfied, but the requirement of the non-negativity of the expectation is nonetheless met.

In Theorem 5.3.1 we obtain a result similar to that of Theorem 5.2.1, but need an additional Condition (5.18) on the tail of the counting variable τ.

Theorem 5.3.1. *Assume that a counting random variable τ is independent of $\{\xi_n\}$. Let $F \in \mathcal{S}^*$. If $\mathbb{E}(\xi) \ge 0$ and there exists $c > \mathbb{E}(\xi)$, such that*

$$\mathbb{P}(c\tau > x) = o(\overline{F}(x)), \tag{5.18}$$

as $x \to \infty$, then the asymptotics (5.6) again hold.

Remark. Condition (5.18) implies that the tail probability of the random variable $c\tau$ decreases faster than that of the increments ξ. That is, $\mathbb{P}(\xi > x)$ asymptotically dominates $\mathbb{P}(c\tau > x)$. It is therefore more likely that ξ exceeds x than it is for $c\tau$ to exceed x, when x is large. Since c is a positive number, it follows that it is also more likely that ξ exceeds x than it is for τ to exceed x (at least when $c \ge 1$). We interpret the Condition (5.18) so that the increment ξ dominates the magnitude of the sum.

Proof. We will present the idea of the proof. The details can be found in [Denisov et al., 2010], proof of Theorem 1, part ii).

Suppose that $\mathbb{E}(\xi) \geq 0$. Clearly $S_\tau \leq M_\tau$, as M_τ is the maximum of S_i's, $0 \leq i \leq \tau$. The proof of inequality (5.10) was done before presenting any conditions for the sign of the expected value of ξ, so it holds in particular when we assume that the expected value of ξ is non-negative. Therefore it is enough to show that $\mathbb{P}(M_\tau > x) \sim \mathbb{E}(\tau)\overline{F}(x)$ holds.

We start by writing $\mathbb{P}(M_\tau > x)$ as a sum of three components. Indeed, for any $N \in \mathbb{N}$ it holds that,

$$\begin{aligned}\mathbb{P}(M_\tau > x) &= \mathbb{P}(M_\tau > x, \tau \leq N) + \mathbb{P}(M_\tau > x, \tau \in (N, \tfrac{x}{c}]) + \mathbb{P}(M_\tau > x, c\tau > x) \\ &= P_1 + P_2 + P_3.\end{aligned}$$

For P_1 we obtain asymptotic equivalence

$$P_1 \sim \mathbb{E}(\tau)\overline{F}(x). \tag{5.19}$$

Both for P_2 and P_3 we have that they are less or equal to something that is $o(\overline{F}(x))$. Indeed, for P_2 we have

$$P_2 \leq KK_1 \int_{N(x)\mathbb{E}(\xi)}^{\frac{bx}{c}} \overline{F}(x-y)\overline{F}(y)\mathrm{d}y = o(\overline{F}(x)), \tag{5.20}$$

as $x \to \infty$. For P_3 we have

$$P_3 \leq \mathbb{P}(c\tau > x) = o(\overline{F}(x)), \tag{5.21}$$

as $x \to \infty$.

Relations in (5.19), (5.20) and (5.21) yield the desired result that $\mathbb{P}(M_\tau > x) \sim \mathbb{E}(\tau)\overline{F}(x)$. □

Before presenting the "complement" of Theorem 5.3.1, we define dominated varying distributions along with intermediate regularly varying distributions.

Definition 5.3.2. A distribution F is called *dominated varying* if there exists a constant c, such that

$$\overline{F}(x) \leq c\overline{F}(2x),$$

for all x.

It is known that any long-tailed and dominated varying distribution with finite expectation belongs to the class $\mathcal{S}^*$.

Definition 5.3.3. A distribution G is *intermediate regularly varying* at infinity, if

$$\lim_{\varepsilon\to 0^+} \limsup_{x\to\infty} \frac{\overline{G}((1-\varepsilon)x)}{\overline{G}(x)} = 1. \tag{5.22}$$

Any intermediate regularly varying distribution is long-tailed and dominated varying. Provided that their expectation is finite, intermediate regularly varying distributions belong to the strong subexponential class.

We are now in a position to state the last theorem concerning the asymptotics of the compound variable. Theorem 5.3.4 is Theorem 7 from [Denisov et al., 2010].

Theorem 5.3.4. *Assume that a counting random variable τ is independent of $\{\xi_n\}$. Let $F \in \mathcal{S}^*$, $\mathbb{E}(\xi) > 0$ and*

$$\overline{F}(x) = O(\mathbb{P}(\tau > x)), \tag{5.23}$$

as $x \to \infty$.
If the distribution of τ is intermediate regularly varying, then

$$\mathbb{P}(S_\tau > x) \sim \mathbb{P}(M_\tau > x) \sim \mathbb{E}(\tau)\overline{F}(x) + \mathbb{P}(\tau > \tfrac{x}{\mathbb{E}(\xi)}),$$

as $x \to \infty$.

Proof. The proof can be found in [Denisov et al., 2010, Section 5]. □

Theorem 5.3.4 does not assume Condition (5.18) to hold. Such a situation would be arising in the case of branching processes. In [Denisov et al., 2010, Section 6], the interested reader can find an application of Theorem 5.3.4 to the branching processes.

Condition (5.23) insinuates that the tail of τ, that is $\mathbb{P}(\tau > x)$, is comparable with or heavier than that of ξ, which is $\overline{F}(x)$. When the tail of τ is heavier than that of ξ, it means that very large values of τ are more likely than the corresponding values of ξ. The interpretation is that the counting (or number of claims) variable dominates the magnitude of the sum. Theorems 5.3.1 and 5.3.4 portray two versions of asymptotics of the tail of a compound variable whose increments have non-negative (or strictly positive) expectation.

Chapter 6

Conclusions and discussion

Wald's identity lets us calculate the expectation of a sum variable that is not necessarily a compound variable. Therefore it extends our ability to determine expectations beyond just the compound variables. Since Wald's identity can be applied to compound variables as well, clearly it should give the same expectation for compound variables as a straightforward calculation of the expectation would. What is interesting though is that the same result can be applied for the case where the counting variable is not independent of the increments as it is in the case of compound variables. We still obtain a similar result for the expectation. So, from the expectations point of view, it does not matter whether the counting variable is independent of the increments, or if a certain type of dependence is allowed.

In Chapter 5 we started to look at the asymptotic tail probabilities of compound variables. A compound variable S_τ could have very large values due to two scenarios. The number of summands might be relatively low, but amongst the summands, there is one that is particularly large. Another situation where the compound variable could get very large values is when the values of summands are moderate or even on the smaller side, but the number of summands is extremely large.

Inspired by these observations we derived asymptotic equivalences in three situations. First we derived somewhat intuitive asymptotics for the situation where the increments have a negative expectation. We also have similar asymptotics for when the increment variable dominates the sum, and the expectation of the increments is non-negative. To pair this result up, we obtained asymptotics in the case where the increments have a strictly positive expectation, and the counting random variable determines the magnitude of the sum. These asymptotic results are of interest to anyone who wants to evaluate the tail probabilities of compound variables, as an insurance company for instance would, if they model the total claim amount as a compound variable.

Bibliography

[Daykin et al., 1994] Daykin, C. D., Pentikäinen, T., and Pesonen, M. (1994). *Practical Risk Theory for Actuaries*. Chapman & Hall.

[Denisov et al., 2010] Denisov, D., Foss, S., and Korshunov, D. (2010). Asymptotics of randomly stopped sums in the presence of heavy tails. *Bernoulli*, 16(4):971–994.

[Embrechts et al., 2008] Embrechts, P., Klüppelberg, C., and Mikosch, T. (2008). *Modelling extremal events: for Insurance and Finance*. Springer.

[Foss et al., 2011] Foss, S., Korshunov, D., and Zachary, S. (2011). *An introduction to heavy-tailed and subexponential distributions*. Springer.

[Foss and Zachary, 2003] Foss, S. and Zachary, S. (2003). The maximum on a random time interval of a random walk with long-tailed increments and negative drift. *Ann. Appl. Probab.*, 13(1):37–53.

[Gut, 2005] Gut, A. (2005). *Probability: A Graduate Course*. Springer.

[Izyurov, 2019] Izyurov, K. (2019). Probability theory - lecture notes.

[Lehtomaa, 2019] Lehtomaa, J. (2019). Risk theory - lecture notes.

[Tuominen, 2010] Tuominen, P. (2010). *Todennäköisyyslaskenta I*. Limes ry.